Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 6 Issue 3, March - 2019

Solving the Internet's Congestion Problem

Assistant Prof. Dr. Omer M. Salih Abudabous
Sirte University Faculty of Sciences –Computer Dep.
omeryomery@yahoo.com

Assistant Prof. Dr. Mosbah M. Alsaade
Sirte University Faculty of Sciences –Computer Dep.
Mosbah_su@hotmail.com

***Abstract*— Congestion, in the context of networks, refers to a network state where a node or link carries so much data that it may deteriorate network service quality, resulting in queuing delay, frame or data packet loss and the blocking of new connections. In a congested network, response time slows with reduced network throughput. In the network layer, when the number of packets sent to the network is greater than the number of packet the network can handle (capacity of network), a problem occurs that is known as congestion, that means congestion occurs when bandwidth is insufficient and network data traffic exceeds capacity. Data packet loss from congestion is partially countered by aggressive network protocol retransmission, which maintains a network congestion state after reducing the initial data load. This can create two stable states under the same data traffic load - one dealing with the initial load and the other maintaining reduced network throughput [1].**

Keywords—Network; Congestion Network; Congestion Control; ECN; TCP.

I. INTRODUCTION

Congestion has been described as an important effect of limited network resources, especially router processing time and link throughput. Traffic directing processes, performed by routers on the Internet and other networks, use a microprocessor. Cumulative router processing time greatly impacts network congestion. In fact, intermediate routers may actually discard data packets when they exceed its handling capability. When this occurs, additional data packets may be sent to make up for unreceived packets, which exacerbates the problem. Network congestion often leads to congestion collapse.[2]

Avoiding network congestion and collapse requires two major components:

- Routers capable of reordering or dropping data packets when received rates reach critical levels
- Flow control mechanisms that respond appropriately when data flow rates reach critical levels

On the other hand, if we choose to allow the traffic, you then have a few choices for how to deal with the congestion. In networking, congestion occurs on shared networks when multiple users contend for access to the same resources (bandwidth, buffers and queues). When the number of packet sent into the network is within the limits, almost all packets are delivered, however the traffic load increases beyond the network capacity. As a result the system starts discarding packets, because routers receive packets faster than they can forward them, one of these two things may happen in the case of congestion. The Sub net may prevent additional packets from entering the congested region until those already present can be processed or the congested routers can discard queued packets to make room for those that are arriving currently[3]. Congestion control refers to the network mechanism and techniques used to control congestion and keep the load below the networks capacity. Congestion handling can be divided into the following:

- Congestion Recovery: Restore the operating state of the network when demand exceeds capacity.

- Congestion Avoidance: Anticipate congestion and avoid it so that congestion never occurs.

Some basic techniques to manage congestion are:-

- End system flow control (ESFC): This is not a congestion control scheme. It is way of preventing the sender from overrunning the buffers of the receiver.

- Network congestion control (NCC): In this scheme, end systems adjust back in order to avoid congesting the network. The mechanism is similar to end-to-end flow control, but the meaning is to reduce congestion in the network, not at the receivers end.

- Resource allocation (RA): This technique involves scheduling the use of physical circuits or other resources, perhaps for a specific period of time. A virtual circuit, built across a series of switches with guaranteed bandwidth is a form of resource allocation. This technique is difficult, but can eliminate network

congestion by blocking traffic that is in excess of network capacity.

II. Congestion Control

When one part of these subnets becomes overloaded, the congestion will be as results [4]. Because routers are receiving packets faster than they can forward them, one of two things must happen:

1. The subnet must prevent additional packets from entering the congested region until those already present can be processed.
2. The congested routers can discard queued packets to make room for those that are arriving.

We now consider the problem of congestion and some possible solutions.

- Preallocation of Resources:
  Preallocation schemes aim to prevent congestion prevent congestion from happening in the first place. For example, we can require that resources be preallocated before any packets can be sent, guaranteeing that resources will be available to process each packet. In virtual circuit networks, for example, the sender opens a connection before sending data. The circuit setup operation selects a path through the subnet, and each router on the path dedicates buffer space and bandwidth to the new circuit.
  The ability of the subnet to reject requests to open connections is an important property of connection oriented networks.
- Traffic Shaping:
  Control the rate at which packets are sent (not just how many). Widely used in ATM networks. At set up, the sender and carrier negotiate a traffic pattern (shape).
  Leaky Bucket Algorithm (LBA) used to control rate in a datagram network. Also in contrast the Token Bucket Algorithm (TBA) causes a token to be generated occasionally, which during idle periods can be saved up.
  Related to traffic shaping is flow specification, where a particular Quality of Service QoS is agreed upon between sender, receiver and carrier.

When the handshake (between sender, receiver) is complete the hosts can start sending data. The data segments will be referred to as packets in the following discussion. The TCP gives a header to every packet. This header contains information about the sender and the packet, to be used by the receiver.

An Internet Service Provider (ISP) can determine how fast it sends traffic over its network. The opposite result of this is the ISP can also slow the rate at which data is moving over its network. This is called artificial congestion. ISP's do this for many reasons, which they claim as network management.

What are affecting this are the established peering agreements. Also, a content provider can be sending begin more traffic than the ISP would like. The larger implications of this are can be an infringement on net neutrality. Network congestion can present itself in many forms. By addressing the causes of network congestion, we can to improve your network.

III. Protocols have affect in congestion positive and negative:

- Border Gateway Protocol (BGP): BGP can be causing network congestion. BGP sends all traffic through the shortest logical path. There is no consideration for how much traffic is already going over that path. No consideration for current data going can result in transit paths becoming overloaded. This overload will create slower speeds, which is network congestion.[10][11]
- Rogue Adapter Broadcasts (RAB): RAB are any foreign devices on your network. This can be as simple as a neighbor coming onto a residential Wi Fi connection or, as severe as a hacker breaking into an enterprise network. That occurs is the rogue adapter finds an entry point, which is usually an error in the network. Then once on the network, they begin to access the Internet. Having an extra device on a network can cause unexpected slowdowns. Besides slowing the network, the bigger problem is the security threat. Any foreign device on a network can become malicious in intent.
- User datagram protocol (UDP): As UDP traffic increases potential problems arise. Unlike TCP, UDP has no mechanism for congestion control leading to wasted bandwidth and poor performance for other competing protocols. UDP uses a simple connectionless communication model with a minimum of protocol mechanism. UDP provides checksums for data integrity, and port numbers for addressing different functions at the source and destination of the datagram. It has no handshaking dialogues, and thus exposes the user's program to

any unreliability of the underlying network; there is no guarantee of delivery, ordering, or duplicate protection. If error-correction facilities are needed at the network interface level, an application may use Transmission Control Protocol (TCP) or Stream Control Transmission Protocol (SCTP) which are designed for this purpose. UDP is suitable for purposes where error checking and correction are either not necessary or are performed in the application; UDP avoids the overhead of such processing in the protocol stack. Time-sensitive applications often use UDP because dropping packets is preferable to waiting for packets delayed due to retransmission, which may not be an option in a real-time system.[12]

- Adding Retransmitting Hubs: When building out a network, there needs to be the integration of hubs. Hubs retransmit data over a network. In an enterprise network, a hub is what connects the network to the public Internet. This connection point offers a prime location for potential congestion. Thus, consider how to integrate the hub within the network.
- Multicasting: Multicasting is where a network allows many computers to speak to each other simultaneously. This is the opposite of Unicast. Unicast is sending traffic to a specific router associated with a specific server. Multicasting is designed to end network congestion. The reality is it could be creating it. As with most things, there are unintended consequences. In multicasting, two packets transferred at the same time can cause a collision. This collision causes network congestion.
- Outdated Hardware: Data transmitted through outdated switches, routers, servers, and Internet exchanges can cause bottlenecks. If the hardware is not optimal, this creates a bottleneck for the transmission of data. The result is network congestion.
- Low Bandwidth: Bandwidth refers to the “size of the pipe” in which Internet data can travel through. If the pipe is not large enough for all the traffic to move through at once, there becomes congestion. This occurs during peak TV streaming hours when Netflix is consuming 40% of the Internet. The result is congestion, as many people are trying to consume large file size streaming.
- Broadcast Storms: A broadcast storm is a situation where there are unexpectedly too many requests on a network. This creates a situation where a network does not have the ability to process all the requests at once. A broadcast storm can be a busy day for e.Commerce or Black Friday sales. Also, a video going viral can cause a similar situation.
- Backpressure Routing: As the throughput of a network reaches capacity, rather than continuing to send data over that network route, you can choose a different path. This is the concept of backpressure routing. Backpressure routing is an algorithmic implementation on a network. It specifies that when a network route begins queuing, traffic is routed over a different path. This solution applies primarily to multi-hop routing, but is effective in eliminating congestion. Although backpressure routing is primarily theoretical, there is increasing potential for its use. As shortest logical path routing of border gateway protocol (BGP) needs to evolve, this is a logical approach to addressing congestion.
- Reconfigure TCP/IP Settings: As traffic is moving over a network, it could be that a sending computer is transferring files faster than the receiving computer can process. The issue with this is that in an un-congested network, packets transfer fast. When they get to the receiving computer, they will become congested. The result is packet loss, as data is not processed. To solve for this issue, you can adjust the TCP/IP settings to slow the request of packets. This can be useful when more computers request on a network. By slowing requests, the receiving computer will be able to manage processing the packets. This can minimize the occurrence of congestion.
- Use a Content Delivery Network: The use of a content delivery network (CDN) has many advantages. Most of the advantages pertain to the global distribution of static content. As about 2/3 of a company’s content is static, this can provide network management benefits. The CDN will place content on edge serves. The result will be less requests coming into your network. If you have segmented your network, this will especially hold true. Less requests will mean less opportunity for congestion. Also, a CDN will assist with bandwidth management. Implementing a CDN can reduce network congestion by placing more requests on edge servers.
- Network Segmentation: Segmenting a network is a process in computer networking. The process is to divide the network into smaller sub-networks. The benefit of segmenting the network is to group assets and groups into specific areas. This will allow for monitoring traffic in groups. It provides grain level insight into the function of the network.

By segmenting, can now reduce network congestion in specific areas of the network. Not having to guess where the congestion is occurring.

- Monitor Your Network Traffic: Monitoring network traffic provides insight about where possible congestion may lie. What this means is that you can make network adjustments to problem areas. The only way to understand if slow network speeds are caused by congestion is to monitor. Network monitoring is how you begin reducing network congestion. You have to understand the distribution of network traffic to analyze reduction in congestion. Once you analyze traffic, then you can provide optimization solutions.
- Choke Packet: To prevent congestion from escalating, the use of a choke packet can be a good strategy. A choke packet is used in network maintenance to prevent the congestion of a network. As a network begins to slow and become congested, a choke packet is sent to slow the output of the sending computer. Decreasing the sending rate is what will allow the receiving computer and routers to catch up. This can prevent the congestion from getting worse and leading to packet loss or a time out.
- Implicit Congestion Notification: Implicit congestion notification is a notification process that is performed at each hop of data transmission. This means there is a node in the sending data packets. The node will pick up information to determine whether there is a possibility of congestion. The benefit to this notification system is there is no extra control messages needed. This is in contrast to our next notification setting, which is explicit congestion notification.
- Explicit congestion notification (ECN): ECN is a notification mechanism that alerts if there is congestion within a network. It works such that there are no packets dropped as congestion begins to occur. In a system where ECN is not used, the congestion notification is dropped packets. This is not an ideal situation as you begin to lose data. The result is packet loss, which can lead to jitter and more congestion. Having a signaling mechanism can allow changes before congestion occurs.[13]
- Prioritize Network Traffic: Preventing congestion can be changing a router setting. By understanding quality of service (QoS) principles, the prioritization of traffic over your network can be important. This is most often used in voice over IP (VoIP) settings. In this situation, VoIP always gets the priority over a network. Prioritization of traffic is ideal for bandwidth utilization. It can ensure there is no one application that is a bandwidth-hog. Also, by adjusting router settings, you can reduce congestion before it begins.
- Use Network Redundancy: Redundancy ensures network availability. What redundancy does is ensure that if one router or network route becomes congested, a second route is used in its place. This would make sure there is no packet loss or a time out due to congestion. The redundancy system would be the failover option for when congestion gets so bad there is a time out. Also, redundancy is used most often when there is a network outage. It prevents an enterprise network form going down. There are many network congestion solutions and I realize I said I was going to provide 10 ways to reduce traffic congestion, but I want to throw in a bonus. But how can router measure congestion? A router might estimate the level of congestion by measuring the percentage of buffers in use, line utilizations, or average queue lengths [6].
- Advantage: Dynamic. Host sends as much data as it wants, the network informs it when it is sending too much.

And the Disadvantage: Difficult to tune. By how much should a host slow down? The answer depends on how much traffic the host is sending, how much of the congestion it is responsible for, and the total capacity of the congested region. Such information is not readily available in practice.

After receiving a choke packet, the sending host should ignore additional choke packets for a short while because packets currently in transmission may generate additional choke packets. How long? Depends on such dynamic network conditions as delay [7].

- Variations exist.
- Overall: Varying methods for congestion control with different levels of effectiveness. More attention being paid to reserving resources so that chances of congestion are reduced and the quality of service (QoS) is more reliable.

IV. TO SOLVE NETWORK CONGESTION

It's a commonly known fact that congestion on the road significantly increases the time the auto vehicles must spend to get to the destination. Exactly the same happens with data networks: congestion considerably slows down the overall network performance. Latency denotes the period of time spent by data on traversing a network segment.[8] High latency, caused by a congested network, slows the speed of the enterprise network performance leading to unsatisfied and angry end users. Here are five ways to reduce the congestion in enterprise networks:

Five ways to reduce the congestion in enterprise networks:

1. Conduct an analysis of the network traffic flows with the help of network monitoring tools. Setup a network sniffer to analyze network traffic, so that underlying troubles in network can be found out and submitted for resolution. While monitoring the network, look into the segments which generate the highest volume of traffic. In case your monitoring system detects that a certain segment generates more traffic than expected, troubleshoot the problem to resolve it. For instance, a computer that floods the network while running a particular program or process may indicate a software networking issue.
2. Network bottlenecks, which are the main reasons of congestion in a network, must be eliminated. Just like a traffic jam caused by a narrowing of a busy four-lane highway to just two lanes, a bottleneck, which is a network's segment unable to handle the amount of traffic coming from its connecting segments, causes intolerable amounts of congestion on a network. Bottlenecks can be eliminated by increasing the segment's bandwidth capacity so it matches the neighboring segments' maximum traffic flow. This can be accomplished by upgrading this network's segment or using different one.
3. TCP/IP protocol settings must be optimized to improve the connection between nodes in your network. You can use TCP tuning techniques to adjust the network congestion avoidance parameters of TCP connections over high-bandwidth, high-latency networks. Sometimes, properly tuned networks can perform up to 5 times better. However, be aware that blindly following instructions without understanding their real consequences can worst the performance as well.
4. Network traffic must be prioritized in compliance with your needs. In some cases congestion can be reduced by prioritizing specific network traffic needs over others. For medium business it can be achieved by reprogramming of network's nodes, such as routers or switches to enable them to identify and give higher priority to the specific types of traffic. For small businesses – enabling the Quality of Service (QoS) feature, which is available in the majority of advanced home or small business networking nodes, will suffice.
5. Finally, to minimize the congestion and correspondingly latency, one must analyze the traffic patterns to find the segment of network, where the congestion occurs, as well as the reason why it happens. After the problem has been outlined one can find an effective solution to avoid congestion and improve the entire infrastructure. On the other hand, if you choose to allow the traffic, you then have a few choices for how to deal with the congestion. Of course, there are pros and cons to each option.
   - Add more bandwidth.
   - Perform quality of service (QoS) on the traffic.
   - Compress the traffic [9].

REFERENCES

[1] MPLS Technology and Application by Bruse Davie & yakkov Rekher: MORGAN KAUFMANN Publishers 2000.
[2] Network Congestion Control "Managing Internet Traffic" by Michael Welzl, Wiley on Communications Networking and Distributed Systems..
[3] Open research issues in Internet congestion control. D Papadimitriou, M Welzl, M Scharf, B Briscoe. 67, 2011. A fault tolerant mechanism for handling permanent and transient failures in a network on chip. M Ali, M Welzl, S Hessler, S Hellebrand. Information Technology, 2007.
[4] M. Allman, V. Paxson, and W. Stevens. TCP Congestion Control, April 1999. RFC 2581.
[5] A.T. Andersen and B.F. Nielsen. A Markovian Approach for Modeling Packet Traffic with Long-Range Dependence. IEEE Journal on Selected Areas in Communications, 16(5), Jun 1998.
[6] D.O. Awduche. MPLS and traffic engineering in IP networks. IEEE Communications Magazine, 37(12):42–47, Dec 1999.

[7] H. Balakrishnan, V.N Padmanabhan, S. Seshan, M. Stemm, and R.H. Katz. TCP behavior of a busy Internet server: analysis and improvements. In IEEE INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings., volume 1, pages 256–262, 1998.

[8] S.C. Borst and D. Mitra. Virtual Partitioning for Robust Resource Sharing: Computational Techniques for Heterogeneous Traffic. IEEE Journal on Selected Areas in Communications, 16(5), Jun 1998.

[9] O.J. Boxma and J.W. Cohen. The M/G/1 Queue with Heavy-Tailed Service Time Distribution. IEEE Journal on Selected Areas in Communications, 16(5), Jun 1998.

[10] Orbit-Computer-Solutions. Com(n.d), Computer Training & CCNA Networking Solutions, Orbit-Computer-Solutions.com, retrieved 8 October 2013,<"Archived copy". Archived from the original on 2013-09-28. Retrieved 2013-10-08.

[11] Sobrinho, João Luís (2003). "Network Routing with Path Vector Protocols: Theory and Applications" (PDF). Retrieved March 16, 2018.

[12] RFC 768 – User Datagram Protocol.

[13] RFC 3168 – Explicit Congestion Notification.