Congestion Policies

Overview

Congestion control and quality of service are the two closely related topics. Improving one means improving.
And ignoring one usually means missing the other. Most technologies used to stop or clear congestion have also enhanced the service standards of the entire network. In these applications, a large amount of knowledge needs to be quickly exchanged over a network of several gigabits per second. For this reason, various types of congestion control used in such broadband networks have been proposed. In this type of congestion control, the idle bandwidth of the network is usually used quickly by increasing the communication speed during communication.
On the other hand, complex types of traffic congestion will occur on the web. Therefore, when these types of congestion control and existing TCP communications coexist, the leading TCP communications will be highly compressed. An essential problem in packet-switched networks is congestion. When the network load-the number of packets sent to the network is greater than the network capacity and the number of packages that the network can handle, network congestion may occur. Congestion control refers to mechanisms and techniques used to regulate congestion and keep the load below capacity. We may wonder why the network is congested. Any system that involves waiting time will have traffic jams. For example, traffic jams occur on highways because any abnormal conditions in the river (such as hourly accidents) will cause blockage. Congestion occurs during the network or internetwork because routers and switches have queue buffers for storing data packets before and after processing. For example, a router has an inbound queue and an outbound queue for each interface.

Congestion factor

- The arrival rate of packets exceeds the capacity of the outbound link.
- There is not enough memory to store the incoming data packet.
- Traffic is heavy.
- The processor is slow.

Slow start Phase

TCP slow start is one of the main steps in the congestion control process. It compares the amount of knowledge that the sender can transmit (the congestion window is known) with the amount of knowledge that the receiver can accept (the receiver's window is known). The lower of the two values is the maximum amount of knowledge that the sender can transmit before receiving an acknowledgment from the receiver.

- The sender tried to talk to the recipient. The sender's initial data packet contains a small congestion window, determining the maximum window to support the sender.
- The recipient confirms the package and responds with his window size. If the receiver does not answer, the sender knows that it can no longer send data.
- After receiving the confirmation, the sender increases the window size of subsequent data packets. The window size gradually increases until the receiver cannot confirm each packet or reaches the sender or receiver window limit.
- After the limit is determined, the slow start task is completed. Other congestion control algorithms will take over the speed of the link.

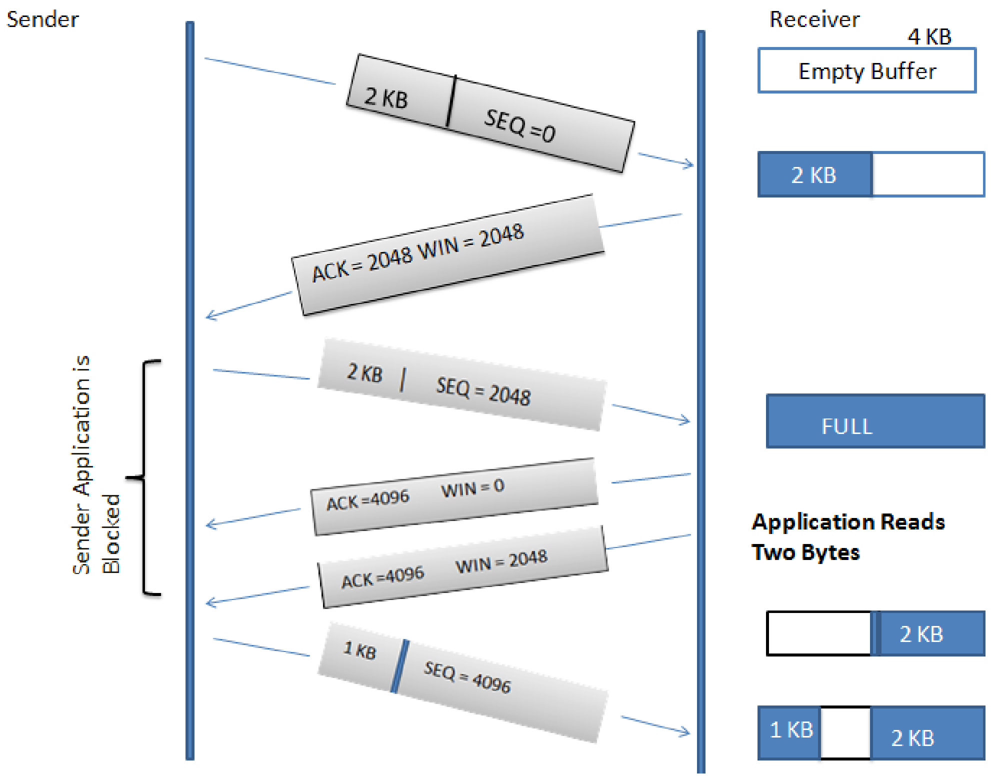

Figure 1: Slow start

Congestion Avoidance Phase

Traditional congestion control schemes help improve performance after congestion occurs. When the load is light, the throughput usually keeps pace with the load. As the load increases, the throughput increases. After the load reaches the network capacity, the throughput will no longer increase. As the load continues to grow, the queue starts to increase, which may cause packets to be dropped. If the load increases more than now, the throughput may drop suddenly, so the network is called congestion. The reaction time initially increases slightly with the load. As the queue is established, the response time increases linearly until the final response time increases sharply because the line restarts due to overflow. The purpose of throughput close to zero is called congestion resolution. This is usually the purpose of the reaction time approaching infinity. The congestion control scheme aims to realize that the network has reached the goal of congestion collapse leading to packet loss and to reduce the load to restore the network to a non-congested state.

Congestion Detection Phase

The transmitter returns to the slow start stage or the traffic jam avoidance stage. When congestion occurs, the size of the congestion window will decrease. The only way the sender can guess that congestion is happening is to retransmit a segment. Retransmission is required to recreate lost packets that are considered discarded by the router due to congestion. Retransmission may occur in two situations: when the RTO timer expires or when three repeated ACKs are received. Before the network can report back information, it must determine its status or load level. In general, the network may also be in one of n possible states. The traffic jam detection function helps assign these states to one of 2 possible load levels: overload or underload (above or below the knee). The Kary version of this function will end with a k-level load indication. For example, the overload detection function can support processor utilization, connection utilization, or queue length.