Multimedia Streaming using TCP or UDP?

Both protocols have advantageous and disadvantageous features. The TCP (Transfer Control Protocol) is understood to be the reliable stream protocol between two end processes and the UDP (User Datagram Protocol) is the unreliable protocol.[1] This means TCP is based on a reliable connection, which entails no lost packets and packets send in order. This advantage comes at the cost of a lot of overhead. Every received packet needs to be acknowledged before the next one is sent. UDP is a connectionless protocol and does not keep track of any lost packets. What is lost is lost and packets may arrive in any order. It is up to the receiving end to sort them out. Because UDP is connectionless the sending end would never know if the receiving end suddenly disconnects, which can lead to problems.[2]

As an example for the UDP multimedia streaming I have chosen the VLC player (VideoLAN Client). It uses UDP among a few other protocols for streaming media over the network. Because UDP cannot normally be used for media streaming, the VLC player uses the RTP (Real-Time Transport Protocol) on top of UDP.[3] The Windows Live Messenger 8.0 is the first of its series to have an inbuilt video calling option and uses the MSNP (Microsoft Notification Protocol) with TCP as basic layer.[4]

The VLC player is a good example of the possibilities of using a UDP based streaming mechanism. UDP itself is not useful as a media streaming protocol, but it allows you to build a new protocol layer on top of it with ease. UDP is very simple and has virtually no inbuilt options except for application multiplexing and checksumming of the header and payload.[5] This means it is up to you to add more functions on an upper layer if needed.

The above image resembles the UDP header format and also shows the simple architecture of UDP. TCP and UDP both assume that IP is used as underlying protocol. The source port is an optional field, but when it is used it indicates the sending port, the destination port is the receiving port, length is the length of the whole datagram and the checksum is the 16-bit one's complement of the one's complement sum of a pseudo header of information from the IP header, the UDP header, and the data, padded with zero octets at the end (if necessary) to make a multiple of two octets.[6] If I were to implement a protocol I would use TCP to start the control and setting up phase, to make sure both parties have all the information they might need. Then I would use a modified UDP, like the VLC player, to deliver the actual multimedia streaming.
References


