Video quality of service

Contents

Chapter five

5. 0 VIDEO QUALITY OF SERVICE

It is easier to understand the QoS issues associated with video traffic on the internet through the understanding of the VoIP. Voice over Internet Protocol (VoIP) seeks to break the difficulties of traditional VoIP and Public Switched Telephone Networks (PSTN) networks through the increased optimization of the network infrastructure and the extension of the customer reach, beyond the PSTN subscribers, without hurting the service offerings and end-to-end capabilities. Its ultimate objective is to reduce the costs of transmission, consolidated network costs, efficient use of equipment and bandwidth as well as greater employee productivity, Gheorghe (2006). The QoS measuring of packet networks that carry real time video and voice considerably differs from exclusive data networks. This is not least because data networks are only concerned with the mitigation of data errors (which are rare) but voice channels are dependent on the perceived or actual quality of the resultant audio output. Owing to the fact that people who use switched-telephone networks do expect great service reliability and high quality of the voice output or high toll quality, it crucial that the internets offerings match the same quality and reliability if they have to compete as equals, Gregori (2002). It is crucial that VoIP systems do not lead to voice quality degradation or cause or be delayed.

The major issue remains the ability of how efficiently can voice be sent through packet networks without hurting QoS. This reflects the urgent need to define a method to carry voice calls across IP networks, complete with the packetization and digitization of voice streams without hurting QoS, while at once describing a networking environment in which data, video and voice transmissions are fully integrated within one system, Ahmad (2001). The merging of telephone signaling, call processing intelligence and packet switching technologies have gone a long way in enabling the consolidation of separate data and voice overlay networks and facilitated the provision of entirely new communication services. The failures of these technologies have in turn driven the emergence of separate data and voice networks, the building blocks for converged data networks that transmit video, data and voice on a single packet network.

5. 1 Video QoS in DiffServ-Aware Multiprotocol Label Switching Network

Jaffar, Hashim, & Hamzah (2009) asseses the quality of Services (QoS) for videos transmitted by (i) Multiprotocol Label Switching (MPLS) that offers greater routing flexibilty, equipment integrationas well as network perfomance and (ii) Differenciated Services (DiffServ), which offers robust and scalable stateless network. By using OPNET and a fixed bandwidth the researches sought to establish the effect of either architecture on the video QoS through the assesment of packet losses, varied video resolutions and delays. The study tested diffServ, MPLS, integrated MPLS and DiffServ as well as Best Effort topologies against a pre-specified background traffic, with three varied services being simulated i. e. AF21, AF11 and Expedited

Forward. Other measures were taken to minimize in herent protoco-specific inneficiencies.

The results indicated that IP Networks just provided FTP, video traffic and http flows for best services using the shortest paths, while longer paths were underutilized, with packets dissipating due to limited network resources. The integrated protocol, which used priority and weighted fair queing, was most efficient, not least because it sent video traffic through the shortest past, while other traffic was routed through the longer path increasing speed and utilization of the network resources. It increases video throughput by 40% than either of the DiffServ and MPLS. It also offered reduced end-to-end delays because of the faster processing rates and PHP network servicing. It also serves FTP and other traffic more efficiently than either protocols by routing video and other traffic separately to avert congestion. Packet losses were markedly higher in MPLS compared to the DiffServ-MPLS protocol to allow for as high as 75% bandwidth aavailability. Effectively, the research results point to the potential of increased use of the DiffServ-MPLS in bolstering video QoS, reduction of packet losses, despite the necessity to better understand other factors such as queuing, traffic policing, congestion avoidance and scheduling in order to ensure absolute efficiency, Bhaniramka, Sun, & Jain (2009).

DiffServ-aware MPLS has been widely adopted modern and future networking technologies (NGN). According to Cho & Okamura (2010), the NGN is increasingly developing the capacity of heterogenous networks convergence, which has made it increasing attractive to the network companies. However, in order to support both technologies, a resource and admision control

(RACF) is necessary for SIP based technology convergence, which are based on the real time, per session services such as those in the vieo and IP telephony. Many research studies of the resource management schemes in the internet architecture have proven effective, and the researchers in Cho & Okamura (2010) proposed the increased use of centralized QoS control scheme that would be employed in engineering traffic in the next generation network core networks that refer to the centralized MPLS-Traffic Engineering. This in line with the findings in Jaffar, Hashim, & Hamzah (2009), which call for the use of compromise systems between MPLS and DiffServ and additionally, both studies recommed the use of multi-level QoS control architectures in order to bolster scalability and simplicity. This is effectively a divide-and-conquer strategy that differentiates interesting objects that athe access networks and the core. The models included in this research reinforce the fact that DiffServ-MPLS bolsters the Quality of Service by reducing packet losses, but moves further in increasing the other congestion avoidance factors, by proposing new generation networks.

Jaffar, Hashim, & Hamzah (2009) calls for greater understanding of other factor that would make DiffServ-MPLS practicable and efficient and Bless & Rohricht (2010) offers such prospects. According to them, the future of telecommunication networks is dependent on the NGN ITU-T framework. The support for end-to-end QoS by providing an efficient control architecture based on the IETFds Next Steps in Signalling frameworks. NGNs are designed in the ITU-T recommendation Y. 2011 (i) and comprises of packed based networks, which offer telecommunication services as well as broadband QoS-enabled transport technologies. Most of the underlying technologies are indepenent in order to facilitate convergence and mobility between mobile and fixed technologies. The new technology responds the deficiencies in the RSVP signaling by using a two layer approach by treating transport signal messages and the signalling applications ae separate, Markopoulou, Tobagi, & Karam (2003). The NSIS framwork is employed to offer overaching end-to-end QoS support and promises greater flexibility and a closer relationship to the actual data path.

5. 2 Data Path Framework/Mechanisms

Other research outcomes have proposed different frameworks to attain end-to-end QoS on the internet especially given the multiple administratice domain. With these mechanisms, the the mechanisms of QoS over the internent fall into two categories i. e. control path and data path. The data path is the basis upon which the the internet QoS is founded and implements actions to be taken by routers regarding the individaul packets to allow the enforcement of different dservice levels, Hentschel, Reinder, & Yiirgwei (2002). The control path mechanisms on the other hand are usually used in the configuration of the network nodes that accord special tratment to packets according to the resource utilization rules.

Packet classifiers determine the packet flow according to the existent rules through signature matching and processing especially for IntServ-enabled routers or by bit-pattern classification. The classification is followed by the measurement of the traffic stream and profile before queiing. Some packets may be dropped if congestion occurs. Similar operations as those performed in Zamora, Jacobs, Eleftheriadis, Chang, & Anastassiou (2000), where specialized programs (SVAs), which would handle similar operations. SVAs are however more adaptive and scale, despite the fact that the they are still technically not as established.

Data path frameworks represent the basic building blocks that facilitate QoS, used in the implementation of the actions that routers must execute on every packet to ensure that varied services are implemented. They are involved with the configuration of the network nodes in relation to which different packets receive differentiated treatment. The basic packet forwarding operation determines the flow of the packet according to the packet classification, which is turn determined by the existent policy and header compression, Chiu, Huang, Lo, Hwang, & Shieh (2003). The data packets are classified according to a general classification or on a bit-pattern classification. General classifications serve transport-level matching of signals based on the packet header tuples, which makes it processing-intensive. The function is indispensable at IntServ-enabled router as well as at the network boundary in the case of DiffServ. Bit-pattern classification ensures that packets are sorted based on a single field in their header. Once they are classified, the packets move onto a logical traffic conditioning, logical instance that has a meter, shaper, marker and a dropper. The markers marks out the differentiated packets, which are subsequently measured by the help of the meter before passing them to a conditioner that compares them to the packet profile, before remarking or dropping out-of-profile packets. The in-profile packets are queued for further processing, which includes reshaping them into the traffic stream, Nisar, Hasbullah, & Said (2009).

5. 3 Queuing Management

The most basic role of QoS is the prevention or minimization of the packet losses, attained through efficient queue management. The packets are lost for two basic reasons i. e. if they are damaged while in transit or dropped because of network congestions. Since packet damage occurs in less than 1% of the case, network congestion becomes the most basic role of QoS, Wang (2001). In order to avoid and control network congestion, a mechanism is necessary both at the intermediate routers as well as the network endpoints and at this time, this is dependent on the TCP Protocol that employs adaptive algorithms (additive increase, slow start as well as multiplicative decrease). Within the routers, the management of queues, in an effort to maximize throughput and reduce delays measured by the network power i. e. ratio of delay to throughput. The buffer space within the network is designed in a way so that it can accommodate short term bursts of data as against continuously holding data. The packets will be dropped if the buffer is full and queue are full, with the newly arriving packets or those that have stayed the longest on the queue being allowed to drop off, Bless & Rohricht (2010).

Locking out packets, when the queues are full can result in a single connection crowding out all other connections on a network. The problem causes a skewed resource utilization system, since a few flows may keep the queue filled up throughout as all other connections are dropped, and in order to avoid these problems, routers must play an active role in the determination of the packets to be dropped. Active queue management allows routers using algorithms such as the Random Early Detection (RED) to control the size of the queue by employing time-based decays to manage the packets that are arrive on a queue, Collins (2001). The possibility of marking a packet is bolstered by the increasing size of the queue since RED employs uses a maximum and minimum queue sizes so that whenever the average queue sizes surpasses the maximum queue size, packets are marked according to the existent policy and possibly dropped according to the algorithm. RED Avoids TCP synchronization through the use of randomization and removal of bias towards bursty traffic.

Martinez, Apostolopoulos, Alfaro, Sanchez, & Duato (2010) deals with the multimedia traffic that has varied requirements and priorities that cannot be met by the best effort only prootcol, nor each of the other individula protocols. With increased internet communication demand, interconnect QoS networks have emerged, spercifffically to deal with video traffic because of its unique charactierisitics comprising of video frames at regular intervals. QoS support for packeted multimedia traffic by technologies such as PCI Express Advanced Switching architecture as well as other technologies included in Jaffar, Hashim, & Hamzah, (2009). Traffic handling policies that are based on deadlines are esily the most effective in scheduling network traffic, whereby packets are labelled according to their respective deadlines, cpupled with random access. This is no different for video traffic, whose packets must be labelled according to their urgency in several ordered queus that are executed concurrently.

The researchers carried out studies similar to Jaffar, Hashim, & Hamzah (2009), in which different conditions were simulated to evaluate the efficiency of the proposed pprotocols for handling traffic. Similarly, three separate approaches were tested by these simulations. The results point towards the support of not only DiffServ-MPLS but also a decline-based policies of scheduling in case of high-speed interconnections, Martinez, Apostolopoulos, Alfaro, Sanchez, & Duato (2010). Specifications such as PCI and InfiniBand were best handled by deadline-based protocols as against traditional protocols, with the performance of the proposed policies being markedly higher than the use of random access buffers. Traffic differentiaon and deadline-based scheduling in any protocol is critical to ensuring efficiency.

5. 4 Scheduling

Controlling the delays on packets is helpful in ultimately ensuring quality of service, and it comprises transmission, propagation as well as queuing, Martinez, Apostolopoulos, Alfaro, Sanchez, & Duato (2010). Packet scheduling defines a mechanism by which packets are selected for transmission from the queued packets, and effectively controls bandwidth allocation to applications, classes and stations. Other than controlling the amount of delay beyond which packets should dropped or the entire operation must be aborted, scheduling also plays a significant role in link sharing. The total bandwidth dedicated to a link may be split between several organizations or protocols, with overloaded links being accorded special attention. Delay guarantees effectively serve as quality guarantees, especially if the process is kept extremely simple. The specialized scheduling algorithms manage the operation of the buffers, and they have two major properties i. e. flow isolation and end-to-end deterministic or guaranteed QoS, Stiller (2009). Flow isolation offers a guarantee that conforming traffic is safe from non-conforming traffic through the partitioning of the network resources, which result in under-utilization of the resources. Isolation ensures fairness from aggressive flows monopolizing the network resources. On the other hand, end-to-end property makes certain that there are statistical and deterministic constraints to the wider network as against simply intermediate nodes.

The scheduling algorithms fall into two broad categories i. e. work-conserving scheduling, which transmits packets for as long as they exist within the buffer and n on-work conserving scheduling that delays transmissions to guarantee delay jitters, while at once presenting the possibility of under-utilization of the resources.

Scheduling can be offered on a per-flow or per-traffic basis, with combinations of the two leading to hierarchical schedules that fall in multiple scheduling disciplines. These include

(i) First Come First Serve, which is the easiest and simplest policy that has neither class or flow differentiation not rate and delay guarantees.

(ii) Priority Scheduling on the other hand offers different queues for every differentiated class of traffic or packets. It is an advanced FCFS discipline that includes priority queues for the priority flows

(iii) Weighted Fair Queuing (WFQ) includes weights and reserved rates for certain links for end-to-end delay guaranteeing on a per-flow basis despite the fact that it is incapable of offering different rate and delay guarantees. This leads into markedly low bandwidth flows being subject to heavy delays, Rosenberg, et al. (2002)

(iv) Earliest Deadline First scheduling discipline on the other hand assigns every packet a deadline within which it must be sent.

There are multiple types of scheduling algorithms, including Process Sharing (DS), De-generalized Process Sharing (GPS), strict priority, Weighted Round Robin (WRR), Weighted Fair Queuing (WFQ) and Earliest Due Date (EDD) among others. FIFO offers only one queue, with each packet being served according to the arrival time. It conserves work, shares the buffering space, but does not guarantee bit rates or packet losses, Zeng (2010). Other subsequent disciplines vary from FIFO because of the capacity to offer differentiated treatment to different data packets. High priority packets are transmitted prior to the low priority packets, and to ensure low priority buffers are not locked out, the scheduling must be carefully implemented.

Strict Priority mechanisms assigns queues a specific priority order, which subsequently determines how the packets are transmitted, effectively allowing for differentiated services in delay and bandwidth allocation, with priority packets enjoying an absolute advantage over regular packets, Gheorghe (2006). This also results in the possibility of aggressive high priority queues completely locking out low priority queues, unless strict priority is coupled by network policing and admission control.

Weight Fair Queue scheduling is founded on the weight ratio of the queues, with the weight (Wi) being assigned to every queue (Qi) that is determined by the existent network policy. In this way, queues are apportioned equal time slots, and so that packets within the same queue are queued separately from those in a different queue so that congestion in one queue only slows that queue, Hentschel, Reinder, & Yiirgwei (2002). This effectively prevents bandwidth abuse, while at once making it possible to offer the necessary delay and bandwidth performance according the allocated bandwidth, leading to mismatches between the delay and bandwidth requirements. Unless high bandwidth is initially allocated for application that require more bandwidth, then there will be inevitable delays for high-bandwidth applications, IXIA (2011).