ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Advanced Network Architectures (M132)

Assignment 3

Quality of Service (QoS) provisioning

Name:	**Evangelos Siatiras**
e-mail:	**EN2190001@di.uoa.gr**

1. QoS and its importance in networking today

Quality of service (QoS) is the capability of providing improved services to a specific network traffic using different technologies like ATM, SONET, and MPLS etc. The main purpose of QoS is to prioritize a specific traffic over another i.e. to take into consideration Jitter, Latency, Packet Loss, and Burst of Jitter and Loss and minimize all these factors for that flow specifically. It should also be considered that prioritizing one traffic flow must not make another fail. Internet provides a best-effort service to all of its applications, i.e., does not make any promises about the Quality of Service (QoS) an application will receive. An application will receive whatever level of performance (e.g., end-end packet delay and loss) that the network is able to provide at that moment. Also note that today's public Internet does not allow delay-sensitive multimedia applications to request any special treatment. All packets are treated equal at the routers, including delay-sensitive audio and video packets. Given that all packets are treated equally, all that's required to ruin the quality of an on-going IP telephone call is enough interfering traffic (i.e., network congestion) to noticeably increase the delay and loss seen by an IP telephone call. Real time applications demand for higher bandwidth and QoS guarantees and to be able to keep the businesses running, researchers are struggling to figure the solution to categorize and implement the routing protocols that separate the CoS (class of service) at the core network. In real time applications like voice and video, QoS is required to set a priority for voice datagram samples and therefore assure packet delivery and reliability.

In practice, when a packet must be forwarded from an interface with queuing, packets requiring low jitter (e.g., VoIP or videoconferencing) are given priority over packets in other queues. Typically, some bandwidth is allocated by default to network control packets (such as Internet Control Message Protocol and routing protocols), while best-effort traffic might simply be given whatever bandwidth is left over and possibly not enough.

2. Scheduling techniques for the provisioning of QoS

The packets belonging to various network flows are multiplexed together and queued for transmission at the output buffers associated with a link. The way queued packets are selected for transmission on the link is known as the link scheduling discipline. The link scheduling discipline plays an important role in providing QoS guarantees. The most important link scheduling disciplines are the following.

- **First-In-First-Out (FIFO)** Packets arriving to the link output queue are queued for transmission if the link is currently busy transmitting another packet. If there is not sufficient buffering space to hold the arriving packet, the queue's packet discarding policy then determines whether the packet will be dropped ("lost") or whether other packets will be removed from the queue to make space for the arriving packet. When a packet is completely transmitted over the outgoing link (i.e., receives service) it is removed from the queue.

- **Priority Queuing** Under priority queuing, packets arriving to the output link are classified into one of two or more priority classes at the output queue. A packet's priority class may depend on an explicit marking that it carries in its packet header (e.g., the value of the Type of Service (ToS) bits in an IPv4 packet), its source or destination IP address, its destination port number, or other criteria. Each priority class typically has its own waiting area (queue). When choosing a packet to transmit, the priority queuing discipline will transmit a packet from the highest priority class that has a non-empty queue (i.e., has packets waiting for transmission). The choice among packets in the same priority class is typically done in a FIFO manner.

- **Round Robin and Weighted Fair Queuing** Under the round robin queuing discipline, packets are again sorted into classes, as with priority queuing. However, rather than there being a strict priority of service among classes, a round robin scheduler alternates service among the classes. In the simplest form of round robin scheduling, a class 1 packet is transmitted, followed by a class 2 packet, followed by a class 1 packet, followed by a class 2 packet, etc. A so-called work-conserving queuing discipline will never allow the link to remain idle whenever there are packets (of any class) queued for transmission. A work-conserving round robin discipline that looks for a packet of a given class but finds none will immediately check the next class in the round robin sequence. A generalized abstraction of round robin queuing that has found considerable use in QoS architectures is the so-called Weighted Fair Queuing (WFQ) discipline. In WFQ the arriving packets are again classified and queued in the appropriate per-class waiting area. As in round robin scheduling, a WFQ scheduler will again serve classes in a circular manner - first serving class 1, then serving class 2, then serving class 3, and then (assuming there are three classes) repeating the service pattern. WFQ is also a work-conserving queuing discipline and thus will immediately move on to the next class in the service sequence upon finding an empty class queue.

3. Models implementing QoS

3.1 Best-effort Service Model

Best effort service model gives no guarantee while delivering the traffic, i.e. it is never known to the sender that either data is delivered or not. Best Effort service model provide performance-oriented service of the network also nodes used are comparatively cheaper than others. As the resources are not allocated so traffic must have to go over the network under traffic load without having the information regarding the packet status whether it's lost, corrupted or delivered. It can be concluded by the various analysis of performance evaluation that best effort service is not best suited for the application which needs specific level of quality of service.

3.2 Integrated Services Model

The integrated service model is used together with resource reservation protocol (RSVP) at every hop to allow per flow-based service state based QoS guarantees [4]. The problem with this model is that it suffers from scalability issues. This model is very complex and if it used for services like providing telephony to subscribers it might result in having wrong billing information.

The three main components of integrated service model are the following.

3.2.1 Classes of Service

- **Guaranteed** It provides the details about the maximum jitter and delay that can occur and also a specific level of bandwidth that is allocated for a certain flow. Guaranteed class is mostly used for those applications that need real time data flow.
- **Controlled Load** It doesn't provide any guarantees of service but provides a constant level to the traffic flow.
- **Best Effort** As it's clear from above that it doesn't provide any assurance about the flow, just does its best for the traffic to reach the specified node. It is mostly used for text based applications.

3.2.2 Control Mechanisms

The traffic is divided into flows based on traffic classes. All the flows are merged within the same class [5]. If there are more than one type of flows in the same router than that number of output queues are used for each line, one for every class.

- **Token bucket filter** It enforces queues and bandwidth in such a way that the maximum limit specified for delay, jitter and bandwidth are met.
- **Random Early Detection** It helps in reducing the dropping of packets when the routers' buffers are full.
- **Weighted Fair Queuing** A queuing management schemes schedules the order in which the queued packets are transmitted.
- **Resource Reservation Protocol** For each traffic flow it reserves the resources like bandwidth and buffer.

3.3 Differentiated Services Model

Differentiated Services Model implements the prioritized model. DiffServ marks packets according to the type of service they desire. In response to these markings, routers and switches use various scheduling strategies to tailor performance to expectations. Differentiated services code point (DSCP) markings use the first 6 bits in the ToS field (now renamed as the DS field) of the IP(v4) packet header. Informally, the packets are coloured so that the forwarding behaviour is indicated. Differentiated Service Model does not focus on individual flow but on the aggregate flow. It provides QoS guarantees [6] for aggregated flow only. It can be used together with Multiprotocol Label Switching (MPLS) and Traffic Engineering (TE). The problem with differentiated services is that it's very poor for end to end QoS guarantees.

4. QoS in MPLS

The data flow of the MPLS comparing with the traditional IP networks which are connectionless is connection oriented and along the pre configured path the packets are forwarded. These pre configured paths are called as LSP's. In order to explain different techniques of implementing QoS in MPLS two different mapping methods are introduced, one is L-LSP (Label-only-inferred-PSC LSP) PSC Protection State Coordination Protocol) is defined per LSP in such a way that the queue behaviour is specified by a certain value such that a flexible DiffServ parameter can be applied per LSP. Similarly E-LSP(EXP-inferred-PSC LSP) where queue behaviour is defined by MPLS exp bit. In MPLS QoS are supported up to eight domains. Now the techniques of QoS for MPLS is the E-LSP (use the EXP bits). This is used by Cisco for frame based MPLS and also reuses the IP code. Another technique is the L-LSP(Cos implicit in the label) where in one Cisco approach for cell based MPLS supports up to four labels of each destination one each for a CoS (Class of Service).Now in case of Cell-mode MPLS QoS the MPLS over QoS options includes Legacy ATM in which PVC (Permanent virtual circuit) acts as a serial link which uses the frame base system such as single PVC mode. Similarly, ATM LSR has two options one single LSP and other Multi VC mode in which along with LSR it will create four LSP per destination. Lower order will contain two bits of EXP bits which will determine CoS (Class of Service) whereas higher order will determine ATM CLP bit. In an IntServ, MPLS support is not described since RSVP label distribution for MPLS is not supported, and capability or out of the reach of IntServ requirement contributing efforts to MPLS network for IntServ reservation. However, MPLS provides full support for DiffServ through small changes in architecture. DiffServ PHB (Per Hop Behavior) and traffic conditioning describe the unchanged for MPLS DiffServ and implemented through LSR, since it performs queuing, marking, policy, metering and shipping etc [7]. MPLS DiffServ can care not only IP traffic but also ATM and frame relay traffic types for the support of Qos with scalability of the network and services. MPLS DiffServ uses PHBs to handle QoS requirements for traffic types with possible extension in network design through increasing LSPs to offer/transport variety of data spectrum. E-LSP (EXP INFARED class LSP) only capable of single transporting and each of them use different encoding mechanisms for marking DiffServ as well as Shim Header [8] implement level stacking for encoding. Shim Header have EXP (3 bits) field to carry 8 service classes for all LSRs, however, E-LSP small number of classes are carried out along with multiple ranking. E-LSPs are defined for admission control by LSRs.

References

[1] G. H. Sabri, "QoS in MPLS and IP Networks," 09 11 2009. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:828354/FULLTEXT01.pdf.

[2] J. F. K. a. K. W. Ross, "Chapter 6," [Online]. Available: https://www.net.t-labs.tu-berlin.de/teaching/computer_networking/.

[3] J. E. I. K. M. A. A. WINARNO SUGENG1, "The Impact of QoS Changes towards Network Performance," 2 2015. [Online]. Available: https://repository.ugm.ac.id/101179/1/p5_3-2.pdf.

[4] X. Y. a. L. D. A. Cheng, "Design and Implementation OF Integrated Services, International Conference on Computer Science and Soft ware Engineering," [Online].

[5] H. B. a. A. H. J. Barakovic, "Multimedia Traffic Analysis of MPLS and non-MPLS Network," [Online].

[6] N. W. Group, "Definition of the Differentiated Services Field," [Online]. Available: http://www.faqs.org/rfcs/rfc2474.html.

[7] N. W. Group, "MPLS Support for DiffServ," [Online]. Available: http://www.faqs.org/rfcs/rfc3270.html.

[8] N. W. Group, "MPLS Label stack encoding," [Online]. Available: http://www.faqs.org/rfcs/rfc3032.html.