SCALABLE VIDEO STREAMING OVER OPENFLOW NETWORKS: AN OPTIMIZATION FRAMEWORK FOR QOS ROUTING

Hilmi E. Egilmez[1,3], Burak Gorkemli[1,2], A. Murat Tekalp[1,3], Seyhan Civanlar[2]

[1]College of Engineering, Koç University, Istanbul, Turkey
[2]ARGELA Technologies, Istanbul, Turkey

ABSTRACT

OpenFlow is a clean-slate Future Internet architecture that decouples control and forwarding layers of routing, which has recently started being deployed throughout the world for research purposes. This paper presents an optimization framework for the OpenFlow controller in order to provide QoS support for scalable video streaming over an OpenFlow network. We pose and solve two optimization problems, where we route the base layer of SVC encoded video as a lossless-QoS flow, while the enhancement layers can be routed either as a lossy-QoS flow or as a best effort flow, respectively. The proposed approach differs from current QoS architectures since we provide dynamic rerouting capability possibly using non-shortest paths for lossless and lossy QoS flows. We show that dynamic rerouting of QoS flows achieves significant improvement on the video's overall PSNR under network congestion.

Index Terms— scalable video, video streaming, OpenFlow networks, optimization, quality of service, QoS

1. INTRODUCTION

Streaming media applications such as videoconferencing, WebTV, and video-on-demand require steady network resources with little or no packet drop and delay variation which can't be always met by the standard best effort Internet. For example, in Scalable Video Coding (SVC), which encodes the video in a base layer and one or more enhancement layers [1], it is crucial that the base layer is streamed without any packet loss or delay variation. Therefore, it is desirable that the network infrastructure supports some means to provide quality of service (QoS) to carry the base layer traffic, while the enhancement layers can be treated as best-effort flows.

Over the past decade, the Internet Engineering Task Force (IETF) has explored several Quality of Service (QoS) architectures, but none has been truly successful and globally implemented. This is because QoS architectures such as IntServ [2] and Diffserv [3] are built on top of current Internet's completely distributed hop-by-hop routing architecture, lacking a broader picture of overall network resources. Although tunneling with MPLS [4] provides a partial solution, it lacks real-time reconfigurability and adaptivity.

OpenFlow is a clean-slate Future Internet architecture, developed by Stanford University [5], which is an open source project aiming at offering a programmable and completely open network to test new Internet concepts such as routing and security that cannot be tested otherwise on the current Internet platforms. We believe that decoupling of the control and forwarding functionalities of routing can be an effective means to provide new QoS architectures over OpenFlow networks.

This paper proposes a new dynamically optimized QoS routing architecture [6] for Scalable Video Streaming over OpenFlow networks. Basics of the OpenFlow network architecture are reviewed in Section 2. Section 3 proposes a new optimization framework for controller design, with two different problem formulations, in order to provide QoS support in OpenFlow networks. The solutions of these problems provide new routing paths for QoS flows which are different than shortest-path best-effort flows. We solve the optimization problems and test the performance of proposed approaches using an open-source network optimization tool, Library for Efficient Modeling and Optimization in Networks (LEMON) [7], led by the COIN-OR project [8]. A simulator, built on top of LEMON, simulates the proposed QoS routing architecture for Scalable Video Streaming. The simulator design and test results are presented in Section 4. Section 5 draws some conclusions.

2. OPENFLOW ROUTING ARCHITECTURE

In the current Internet architecture, the routers perform both routing control and packet forwarding functions. The key difference in OpenFlow is that the forwarding (data) and routing (control) layers are decoupled; where the forwarding function stays within the OpenFlow routers (forwarders) while the routing control function is handled by a separate controller layer, which is the brain of the network (could be centralized and/or possibly federated). Forwarding tables are dynamically uploaded to forwarders by the controller layer. As shown in Fig. 1, the controller layer controls many forwarders. Various studies demonstrate that even a single controller solution is highly scalable. The data path traverses

the forwarders and provides the actual forwarding of the IP packets. OpenFlow brings a unique advantage of designing and programming the controller to generate different forwarding tables for different flows.

With OpenFlow, it is possible to use various routing algorithms (rather than the typical shortest path) within the controller to generate forwarding tables that govern different isolated flows such as the QoS flows in the data plane as illustrated in Fig. 2. This is the key concept that we will be leveraging in this paper.
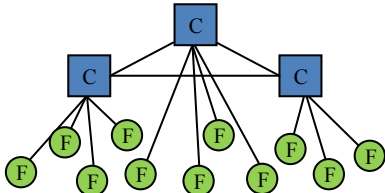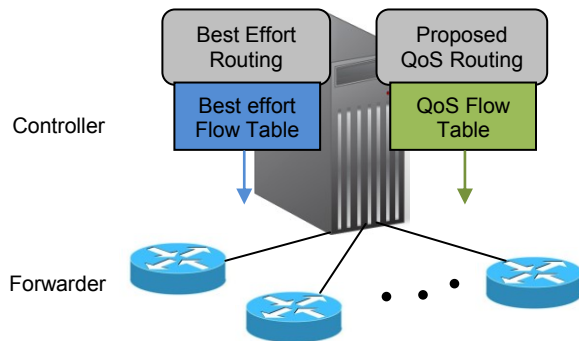

Figure 1: OpenFlow controller and forwarder layers

Figure 2: Controller uploads forwarders' flow tables.

The optimization problems posed in Section 3 dynamically update the optimum route for QoS flows which in turn are translated into flow tables dedicated to QoS flows, without manipulating best-effort flow tables. For this model, some of the basic assumptions are as follows:

- QoS routing should optimize a different cost function than simply the path length. Routes that have larger capacity even with longer distances may be more preferable to shorter routes that may cause packet loss;
- QoS traffic may have preemptive rights based on the traffic priority and importance. Meaning, when QoS traffic is placed on a route, more packet loss may incur on best effort traffic on the shared route;
- QoS routing may select optimal routes based on traffic patterns and packet loss estimates of best effort traffic.

Thus, the routing algorithm that optimizes the QoS traffic must account for the existence of best effort traffic and prevent serious congestion of it. Since the dominating traffic volume in the Internet is best effort, any performance optimization process which cares about QoS sessions cannot ignore the impact on best effort traffic.

3. OPTIMIZATION OF QOS ROUTING

This section poses two optimization problems for QoS flow routing. Before stating the problems in Sections 3.2 and 3.3, we first present the basic definitions and notation.

3.1. Definitions and Notation

The traffic stream in the network is decoupled into three distinct streams:

- SVC base layer video is defined as <u>lossless QoS</u> (i.e., does not tolerate any packet losses) traffic,
- SVC enhancement layers are defined as <u>lossy QoS</u> (i.e., tolerates some packet losses) traffic,
- The rest of the traffic is defined as <u>best-effort</u>.

If there are multiple enhancement layers of a video stream, they are treated as a bundle to be routed together. We can decouple the base layer from enhancement layers and separately route it (by simply using a different port number).

Problem I only reroutes lossless QoS traffic under congestion conditions, while the rest of the traffic remains on their shortest path routes. This implies two flow tables are generated by the controller; one for the QoS flows; and one for the rest of the traffic.

In Problem II, lossless QoS traffic as well as enhancement layers are rerouted, while the best effort traffic remains on its shortest path routes. The lossless QoS and lossy QoS traffic streams can choose different routes. This implies three flow tables are generated by the controller; one for the lossless QoS flows; one for lossy QoS flows; and last one for the best effort traffic. By rerouting the enhancement layers, we further optimize the performance of the received SVC encoded video, as we encounter less packet losses on these layers.

Let each OpenFlow node (forwarder) be controlled by one controller. Pair of nodes are defined by i and k where $i = 1,2,\ldots,N$, $k = 1,2,\ldots,N$ with $i \neq k$ where N is the total number of nodes in the network. Distinct routes between node i and k are denoted as $r_1, r_2 = 1, 2,\ldots, N_{ik}$ where r_1 is the route for lossless QoS traffic (base layer) and r_2 is the route for the lossy QoS traffic (bundle of enhancement layers), and N_{ik} is the total number of distinct routes between nodes i and k. We use L_{ik}^r for length and C_{ik}^r for capacity of a route r between nodes i and k. Length of a route can be the number of hops or the propagation delay from source to destination. Capacity can be a measure of bandwidth in bps. Moreover, in our formulations, we have used three separate traffic variables. $Q_{ik}^r(t)$, $E_{ik}^r(t)$ and $B_{ik}^r(t)$ are the amounts of lossless QoS traffic, lossy QoS traffic and best-effort traffic on route r at time t. In general, we can assume $B_{ik}^r(t) \gg Q_{ik}^r(t) + E_{ik}^r(t)$ since the best-effort traffic dominates.

3.2. Optimization Problem I

Problem I aims to route base layer flows such that they encounter no packet losses at any time. By doing this, enhancement layers and best-effort traffic may encounter more packet losses to enable the base layer to survive.

The cost function of the optimization Problem I, shown in (1) below, is the weighted sum of the route length of rerouted base layer traffic and packet loss rate on the remaining traffic (i.e., the enhancement layers and best-effort traffic) as a result of rerouting the base layer.

$$\underset{r_1}{\text{minimize}} \ (1-\lambda)L_{ik}^{r_1} + \lambda \, PLT_{ik}^{r_1}(t)$$
$$\text{subject to} \ \ L_{ik}^{r_1} < L_{max} \tag{1}$$

where the packet loss rate on any route r at time t, $PLT_{ik}^{r}(t)$, is calculated as:

$$PLT_{ik}^{r}(t) = \begin{cases} 0 & C_{ik}^{r}(t) \geq Q_{ik}^{r}(t) + B_{ik}^{r}(t) + E_{ik}^{r}(t) \\ \dfrac{Q_{ik}^{r}(t) + B_{ik}^{r}(t) + E_{ik}^{r}(t) - C_{ik}^{r}(t)}{B_{ik}^{r}(t) + E_{ik}^{r}(t)} & C_{ik}^{r}(t) < Q_{ik}^{r}(t) + B_{ik}^{r}(t) + E_{ik}^{r}(t) \end{cases}$$

$i =1,2,\ldots, N$, $k =1,2,\ldots, N$, $r_1 = 1,2,\ldots, N_{ik}$ where $0 \leq \lambda \leq 1$ and assuming that there exists an $r = r_1$ such that $C_{ik}^{r_1}(t) > Q_{ik}^{r_1}(t)$ which satisfies zero packet loss for QoS traffic, i.e., $PLQ_{ik}^{r_1} = 0$. The inequality constraint in (1) specifies the maximum tolerable length (delay) of the new QoS route.

Solution of this problem is the minimum cost route r_1 from the feasible set of routes having shorter length than L_{max}. The weight λ determines the relative importance of the new route length and the packet losses. For large λ, route selection would be more sensitive to total packet losses on the QoS route. Vice versa, for small λ our problem converges to shortest-path algorithm. Current best-effort Internet routing applies shortest-path algorithm which only takes the length variable into account. When the shortest-path route is congested, the video stream's performance would be poor due to losses especially for SVC base layer. As a solution to this, we proposed the optimization problem formulation in (1) which guarantees the transmission of SVC base layer without any loss by rerouting of QoS flows on other available non-shortest path routes.

3.3. Optimization Problem II

In Problem II, we define the base layer as lossless and enhancement layers as lossy QoS traffic, so that both base and enhancement layers of the video are rerouted. Our goal is to reroute QoS flows (both base and enhancement layer packets) in such a way that there would be no packet losses in QoS flows corresponding to SVC base layer while packet losses on the links that has the rerouted enhancement layers and the rerouted base layer are minimized.

The formulation for Problem II is as follows:

$$\underset{r_1, r_2}{\text{minimize}} \ (1-\lambda)\left\{L_{ik}^{r_1} + L_{ik}^{r_2}\right\} + \lambda \left\{PLT_{ik}^{r_1}(t) + PLT_{ik}^{r_2}(t)\right\}$$
$$\text{subject to} \ \ L_{ik}^{r_1} < L_{max}, L_{ik}^{r_2} < L_{max} \tag{2}$$

$i =1,2,\ldots, N$, $k =1,2,\ldots, N$, $r_1,r_2 = 1,2,\ldots, N_{ik}$, $0 \leq \lambda \leq 1$ and assuming there exists $r = r_1$ such that $C_{ik}^{r_1}(t) > Q_{ik}^{r_1}(t)$ which satisfies $PLQ_{ik}^{r_1} = 0$.

The key difference between the formulations of Problem I and II is the way the enhancement layers of the SVC video is treated. In Problem I, the enhancement layer is considered as best-effort traffic while in Problem II it is a QoS traffic which may encounter packet drops.

The problems proposed in this work are examples of minimum cost flow optimization problems in network optimization theory. In the literature, there are many algorithms available in order to solve min-cost flow problems [9]. We solved the proposed problems by using the "LARAC algorithm" which is implemented in the network optimization tool LEMON. It is a polynomial-time algorithm [10] having complexity $o\left((m + n\log n)^2\right)$ where n and m are number of nodes and links, respectively.

4. SIMULATIONS

In order to simulate the OpenFlow architecture and test the performances of the proposed routing formulations (see Section 3), we implemented a simulator by using the network optimization library LEMON. LEMON is a C++ template library that provides efficient implementations of optimization algorithms for combinatorial optimization problems with graphs and networks. LEMON also supports efficient data structures such that are useful in creating different network topologies. For simulations, we created the simple network topology shown in Fig.3. For all links, we set length to be 1 (hop-count) and link capacity to be 10Mbps. The traffic (congestion) on each link is modeled as an independent Poisson random variable which is commonly used for Internet traffic modeling for small scale network topologies [11].
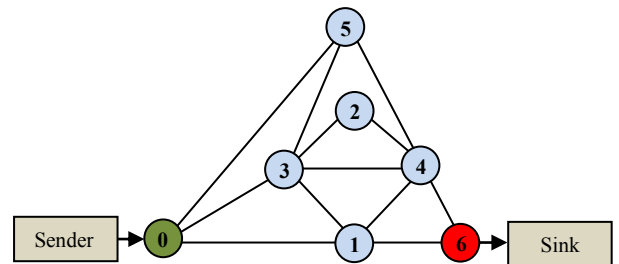

Figure 3: Network topology used in simulations

In the simulation scenario, node 0 is the SVC streaming server and node 6 is the SVC streaming client. The looped MPEG test sequence "Train" with 704 x 576 resolution is used throughout the tests. Looped sequence has 714 frames lasting about 24 seconds. It is encoded by the SVC reference software Joint Scalable Video Mode (JSVM 9.18) to obtain SVC base layer and an enhancement layer. The base layer is encoded at 530 kbps (31.52 dB) and the enhancement layer is encoded at 1833 kbps (39.97 dB). Group of Pictures (GoP) size is set as 16 frames.

The simulator generates new routes for the SVC video stream depending on which optimization problem (I or II) is being solved. The rerouting is performed at each second, which corresponds to approximately 2 GoPs. For each time interval (1 second) the simulator performs the following:

- calculates $PLT_{ik}^{r}(t)$;
- solves the routing optimization problem by using calculated packet loss rate $PLT_{ik}^{r}(t)$;
- reroutes video packets, according to Problem I or II formulation;

The simulator is designed in such a way that we can track which specific video packets are lost. By matching those lost packets with the Network Abstraction Layer (NAL) units of the SVC video stream, we decide which NAL units are lost in the received video stream. Then, the received stream is decoded and the PSNR values are measured. Since we are modeling the congestion with independent Poisson random variables, we have to evaluate the average performances of the two approaches. So, the experiment is repeated 20 times with different best effort traffic loads and the corresponding average PSNR values are calculated.

Three different scenarios are executed, and the results are given in Fig.4 in terms of received video quality by calculating the PSNR values with respect to original raw video. We first simulate the traditional best-effort Internet performance where no route switching is performed, during which SVC video stream follows the shortest path route. In Fig.4, we observe the severe decrease in the received video quality in this scenario. The second and third scenarios are the results corresponding to solution of Problems I and II, respectively. We observe that both Problem I and II solutions outperform best-effort internet, and third scenario gives the best performance.
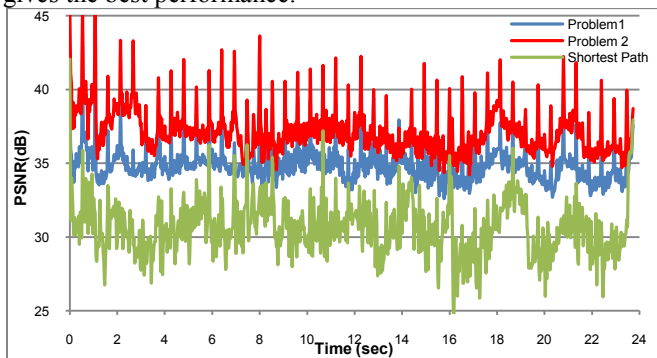

Figure 4: Comparison of routing scenarios

	$\lambda = 0.1$	$\lambda = 0.3$	$\lambda = 0.5$	$\lambda = 0.7$	$\lambda = 0.9$
Problem 1	30.63 dB	31.20 dB	33.21 dB	34.18 dB	34.87 dB
Problem 2	30.63 dB	31.27 dB	33.39 dB	35.86 dB	37.16 dB

Table 1: Effect of λ on video quality in terms of PSNR

In the final tests, the effect of weight λ (see Section 3) on the received video quality in terms of average PSNR is evaluated. Results are shown in Table.1. Note that, for large λ the route selection would be more sensitive to the packet losses, as expected.

5. CONCLUSIONS

We propose and solve different optimization problems to reroute the base layer of SVC video streams losslessly on alternate routes by leveraging the flexibility of the control layer of the OpenFlow architecture. By analyzing two variations of the proposed optimization framework, we have observed that the average quality of video streams is improved by 14% if only the base layer is rerouted. By rerouting the enhancement layer along with the base layer, the quality is further improved by another 6.5%. Thus, we recommend optimizing routes for both the base and enhancement layers using problem formulation II in OpenFlow networks, under congestion while keeping the best effort traffic on its current shortest path route.

6. REFERENCES

[1] H. Sun, A. Vetro, and J. Xin, "An Overview of Scalable Video Streaming," *Wireless Communications and Mobile Computing*, vol. 7, no. 2, pp. 159-172, February 2007.
[2] R. Braden, D. Clark, et.al. "Integrated services in the Internet architecture: an overview," RFC 1633, Jun. 1994.
[3] S. Blake, D. Black, et.al., "An architecture for differentiated services," RFC 2475, Dec. 1998.
[4] E. Rosen, and Y. Rekhter, "BGP/MPLS VPNs" RFC 2547, Mar. 1999.
[5] "OpenFlow." [Online]. Available: http://openflowswitch.org
[6] S. Civanlar, M. Parlakisik, A.M. Tekalp, B. Gorkemli, B. Kaytaz and E. Onem, "A QoS-Enabled OpenFlow Environment for Scalable Video Streaming", *IEEE Globecom 2010 Workshop on Network of the Future (FutureNet-III)*, Miami, USA, 2010.
[7] LEMON, "Library for Efficient Modeling and Optimization in Networks." [Online]. Available: http://lemon.cs.elte.hu
[8] COIN-OR, "Computational Infrastructure for Operations Research." [Online]. Available: http://www.coin-or.org
[9] D.P. Bertsekas, *Network Optimization: Continuous and Discrete Models*, Athena Scientific, Belmont Mass., May. 1998.
[10] Y. Xiao, K. Thulasiraman, and G. Xue, "GEN-LARAC: A Generalized Approach to the Constrained Shortest Path Problem under Multiple Additive Constraints," *Proc. Int'l Symp. Algorithms and Computation (ISAAC),* pp. 92-105, 2005.
[11] J. Cao, W. S. Cleveland, D. Lin, and D. X. Sun, "Internet Traffic Tends Toward Poisson and Independent as the Load Increases," *Nonlinear Estimation and Classification*, eds. C. Holmes, D. Denison, M. Hansen, B. Yu, and B. Mallick, Springer, New York, pp. 83-109, 2002.