End-to-End Transmission Performance Optimization Based Routing Selection Algorithm for Software Defined Networking

Guixiang Jiang(✉), Rong Chai, and Haipeng Li

Key Lab of Mobile Communication Technology,
Chongqing University of Posts and Telecommunications, Chongqing, China
825882445@qq.com, chairong@cqupt.edu.cn, lihaipeng814@qq.com

Abstract. Software-defined networking (SDN) is a new networking paradigm enabling innovation through decoupling control plane from data plane and providing programmability for network application development. Specific research focus has been placed to achieve route optimal selection in SDN scenario. In this paper, we study the problem of route selection for a user flow in a SDN scenario consisting of a number of switches and propose an end-to-end transmission performance optimization based routing selection algorithm. We jointly consider the characteristics of user flow and service capability of the network, and formulate the arrival curve of user flow and the service curve of switches by applying Network Calculus theory. The transmission performance of user flow, defined as effective bandwidth is evaluated and the route offering the maximum effective bandwidth is selected as the optimal route. Numerical results demonstrate the effectiveness of the proposed algorithm.

Keywords: SDN · Routing algorithm · Network calculus
Effective bandwidth

1 Introduction

The rapid growth of data services with different quality of service (QoS) requirements poses challenges to network architecture and management mechanisms. Traditional Internet technologies can hardly meet the increasing service requirements due to tightly coupled control and date plane, inflexible network architecture and complicated network and service management mechanisms. To tackle these challenges, software-defined networking (SDN) technology has been proposed [1].

As an emerging network technology, SDN is one of the most promising approaches to dynamically programming networks and managing user services via building customized solutions which are manageable, dynamic, cost-effective and adaptable. In SDN, controllers in control plane are responsible for controlling and managing data forwarding devices such as switches in a centralized

K. Long et al. (Eds.): 5GWN 2017, LNICST 211, pp. 28–37, 2018.
https://doi.org/10.1007/978-3-319-72823-0_4

manner, thus achieving the separation between data forwarding plane and the control plane. This new network technology permits separating network control functions, e.g., routing from switches so that routing algorithms can be designed in controllers [2,3].

In recent years, the problem of routing algorithm design for SDN has been stressed and several routing algorithms have been proposed. Some references consider applying network function virtualization (NFV) technology to achieve performance enhancement of routing algorithms in SDN. In [4], the authors present a QoS-aware virtualization-enabled routing framework to fulfill the QoS requirement of multiple clients in a SDN. To isolate and prioritize tenants from different clients, a network virtualization algorithm is designed to create a subnet for each tenant, and then a QoS-aware routing algorithm is proposed to maximize the minimum residue bandwidth of links. The authors in [5] present an approach for building multicast mechanisms over SDN. By applying NFV technology, transcoding functions are embedded into some of the servers, switches or routers. A multicast routing algorithm is then proposed to ensure each multicast flow traverses through the nodes embedded NFV technology before reaching the destination, while the total routing cost is minimized at the same time.

User QoS requirements and network characteristics are jointly considered in routing algorithm for SDN. In [6], the authors propose a probabilistic QoS routing mechanism for SDN which jointly considers link probability and bandwidth availability of the network. By applying Bayes theorem and Bayesian network model, link probability is determined based on which the path consisting of the links of the highest probabilities is selected under that satisfies the bandwidth constraint. The authors in [7] stress resource requirements of flows in the core network. Defining resource requirements including transmission bandwidth and flow tables as resource preference, the authors quantize the resource preference of different network flows based on analytic hierarchy process (AHP) method and propose a resource preference aware routing algorithm which matches each flow to the path with the largest preference degree.

In [8], the authors consider transmitting a set of unicast sessions in SDN with each session having a QoS requirement on transmission throughput. Assuming each session is associated with a collection of packet forwarding rules, the authors propose a rule multiplexing scheme, in which the same set of rules is deployed on the common nodes of multiple pathes. Based on the multiplexing scheme, a route selection and rule placement strategy is proposed with the objective of minimizing rule space occupation of all the sessions.

In previous works, the transmission characteristics such as network bandwidth and throughput are taken into account in designing optimal routing strategy, however, the end-to-end transmission performance fails to be stressed, which may result in undesired transmission performance. In this paper, we study the problem of routing algorithm design for one user flow in SDN. Particularly, we consider service preference on high data rate and propose an end-to-end transmission performance optimization based routing algorithm. To quantitatively analyze the end-to-end transmission performance of user flow, we apply Network

Calculus theory, which enables the joint consideration of the service arrival characteristics of the user flow and the serving ability of the intermediate switches, thus achieving comprehensive evaluation of the transmission performance of user flow.

The rest of this paper is organized as follows. Section 2 describes the system model considered in this paper. Section 3 discusses the optimal route selection algorithm. In Sect. 4, we examine the end-to-end transmission performance of user flow. Simulation results are presented in Sect. 5. Finally, Sect. 6 concludes the paper.

2 System Model

In this paper, we consider a SDN model as shown in Fig. 1. The system consists of one SDN controller and a number of switches. We assume that a user flow needs to be transmitted from a source switch (SS) to a destination switch (DS). In the case that no direct transmission link between SS and DS exists, route selection strategy should be designed for the user flow.

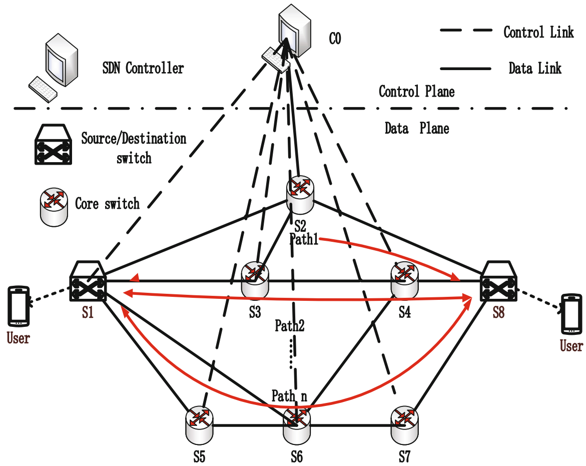

Fig. 1. System model

We assume in the SDN scenario considered, there exists M_0 switches in addition to SS and DS. For convenience, we denote Φ as the set containing all the switches. To characterize the connection status between switches, we introduce binary connection identifiers. Let $\delta_{i,j}$ denote the connection identifier between switch i and switch j, i.e., $\delta_{i,j} = 1$ if switch i is an adjacent node of switch j, thus there exists an available link between switch i and switch j; otherwise, $\delta_{i,j} = 0$, $i \in \Phi$, $j \in \Phi, i \neq j$. In this paper, we assume that the network topology is given, hence, $\delta_{i,j}$ are known constants.

3 Optimal Route Selection Algorithm Description

In this paper, we assume that the user flow has a minimum data rate requirement, to reduce the computation complexity of route selection, we first propose a candidate route selection scheme based on user rate requirement and network characteristics, and then an optimal route selection is proposed to select the optimal route among all the candidate routes.

3.1 Candidate Route Selection Scheme

Given a SDN scenario, we first search for all the possible routes connecting SS and DS. We assume that the number of possible routes connecting SS and DS is M, and denote the mth route as L_m. For convenience, we denote the switches along L_m as $\mathrm{S}_{m,0}$, $\mathrm{S}_{m,1}$, $\cdots$, S_{m,N_m}, where N_m denotes the link number of L_m. We can obtain that L_m is an available route connecting SS and DS provided that $\mathrm{S}_{m,0}$ and S_{m,N_m} are respectively SS and DS and the following condition meets. For simplicity, we assume that the jth switch of L_m is labeled as the m_jth switch of the network, $1 \leq j \leq N_m$, we can achieve that the m_{j-1}th switch and the m_jth switch should be adjacent nodes of the network, i.e.,

$$\delta_{m_{j-1},m_j} = 1. \tag{1}$$

Denote the data rate of L_m as R_m, $1 \leq m \leq M$, we define R_m^0 as the minimum data rate of all the links of L_m, i.e.,

$$R_m^0 = \min\{R_{m,j}^0\},\ 1 \leq j \leq N_m \tag{2}$$

where $R_{m,j}^0$ denotes the data rate of the jth link of L_m. In this paper, we assume that the link capacity is large enough, thus the data rate of one link is determined by the port rate of the two switches being connected by the link. We denote the import and outport rate of $\mathrm{S}_{m,j}$ as $R_{m,j}^{\mathrm{in}}$ and $R_{m,j}^{\mathrm{out}}$, respectively, $0 \leq j \leq N_m$. The data rate of the jth link of L_m can be defined as the smaller one of the outport rate of $\mathrm{S}_{m,j-1}$ and the import rate of $\mathrm{S}_{m,j}$, $1 \leq j \leq N_m$, which can be expressed as:

$$R_{m,j}^0 = \min\{R_{m,j-1}^{\mathrm{out}}, R_{m,j}^{\mathrm{in}}\},\ 1 \leq j \leq N_m \tag{3}$$

Assuming user flow has a minimum data rate requirement, denoted by R^{min}, only the routes meeting this constraint can be selected as the transmission route. Denoting Ψ as the set of the candidate routes, we obtain:

$$\Psi = \{L_m | R_m^0 \geq R^{\mathrm{min}},\ 1 \leq m \leq M\}. \tag{4}$$

3.2 Optimal Route Selection Strategy

To obtain the optimal route, we examine the transmission performance of all the candidate routes, and select the one offering the best performance. In particular, we focus on the end-to-end transmission performance and evaluate the throughput of the candidate routes by jointly considering the characteristics of user flow and the service capability of the switches. Denote T_m, $1 \le m \le M$, as the throughput of the mth route, we obtain the optimal route, denoted by L_{m^*} as:

$$L_{m^*} = \text{argmax } \{T_m\}. \tag{5}$$

The detail evaluation of T_m will be discussed in the following section.

4 End-to-End Transmission Performance Evaluation

In this section, the transmission performance of the candidate routes between SS and DS is evaluated based on Network Calculus theory [9], which is an efficient tool to analyze flow control problems in networks mainly from a lower bound (i.e., worst case) perspective. The essential idea of network calculus is to transform complex non-linear network systems into analytically tractable linear systems by using alternate algebras, i.e., the min-plus algebra and max-plus algebra.

4.1 Brief Introduction of Network Calculus Theory

In this subsection, some concepts defined in the theory of network calculus will be introduced briefly.

Definition 1: Min-plus Convolution. Let f and g be two functions or sequences of time, the min-plus convolution of f and g is defined as follows:

$$(f \otimes g)(t) = \begin{cases} \inf\limits_{0 \le s \le t} \{f(t-s) + g(s)\}, & t > 0 \\ 0, & t < 0. \end{cases} \tag{6}$$

Definition 2: Arrival Curve. Given a wide-sense increasing function $\alpha(t)$ defined for $t \ge 0$, a flow $F(t)$ is constrained by $\alpha(t)$ if for all $s \le t$, following inequality holds:

$$F(t) - F(s) \le \alpha(t-s), \tag{7}$$

we say that $\alpha(t)$ is the arrival curve of $F(t)$.

Definition 3: Service Curve. Consider a system S and a flow with the input and output functions respectively being $F(t)$ and $O(t)$ passing through S, we say that S offers a service curve $\beta(t)$ to the flow if and only if $\beta(t)$ meets following conditions:

$$\beta(0) = 0, \tag{8}$$

$$O(t) \ge (F \otimes \beta)(t). \tag{9}$$

Definition 4: Effective Bandwidth. Consider a flow with arrival curve $\alpha(t)$ passing through a system S, for a given delay $L = \theta + \frac{s_b}{r}$, where θ and s_b denote respectively the service delay and the burst tolerance of S, and r denotes the service rate of S, we define the effective bandwidth of the flow, denoted by $e_L(\alpha)$, as the required throughput at t to serve the flow in a work conserving manner, i.e.:

$$e_L(\alpha) = \sup_{s \geq 0}\{\frac{\alpha(s)}{s+L}\}. \tag{10}$$

From the definition of effective bandwidth, it can be seen that effective bandwidth represents the maximum transmitted data packets for a given time interval, thus is equivalent to transmission throughput, which can be applied to characterize the end-to-end transmission performance of the mth candidate route, i.e., T_m in (5), when the flow with arrival function being $\alpha(t)$ passes through a system with the service curve being $\beta(t)$. In the following subsections, we will examine the transmission performance of candidate routes based on Network Calculus theory. The arrival curve of user flow and the service curve of switches are formulated, then the effective bandwidth of each route is evaluated.

4.2 Modeling Arrival Curve of User Flow

In this paper, referring to [10], we characterize the arrival curve of user flow as traffic-specification (T-SPEC) model [11]. We define the set of T-SPEC parameters $(r^{(\mathrm{m})}, s^{(\mathrm{m})}, r^{(\mathrm{s})}, s^{(\mathrm{b})})$, where $r^{(\mathrm{m})}$ and $s^{(\mathrm{m})}$ represent the maximum arrival rate and the maximum packet size, respectively, $r^{(\mathrm{s})}$ and $s^{(\mathrm{b})}$ represent the sustainable arrival rate and the burst tolerance, respectively, the arrival curve of the user flow can be modeled as:

$$\alpha(t) = \min\{r^{(\mathrm{m})}t + s^{(\mathrm{m})}, r^{(\mathrm{s})}t + s^{(\mathrm{b})}\}. \tag{11}$$

4.3 Modeling Service Curve of Switches

In this paper, we assume that the switches are latency-rate (LR) system, of which the service curve can be characterized by the service rate and the transmission latency. For the jth switch of L_m, the service curve can be modeled as:

$$\beta_{m,j}(t) = R_{m,j}(t - \theta_{m,j})^+ \tag{12}$$

where $R_{m,j}$ and $\theta_{m,j}$ are respectively the service rate and the latency of the jth switch of L_m, $(x)^+ = \max\{x, 0\}$. According to Network Calculus theory, the joint service curve of L_m can be expressed as:

$$\beta_m(t) = \beta_{m,1}(t) \otimes \beta_{m,2}(t) \cdots \otimes \beta_{m,N_m}(t) = R_m(t - \theta_m)^+ \tag{13}$$

where $R_m = \min\{R_{m,1}, R_{m,2}, \cdots, R_{m,N_m}\}$, $\theta_m = \sum_{j=1}^{N_m} \theta_{m,j}$.

4.4 Effective Bandwidth of the Candidate Routes

Substituting $\alpha(t)$ obtained in (11) into (10), and calculating L based on $\beta_m(t)$ obtained in (13), after some simple mathematical manipulations, we can obtain the effective bandwidth when the user flow transmitting through L_m, i.e.,

$$T_m = \max\left\{r^{(\mathrm{s})}, \frac{s^{(\mathrm{m})}}{\theta_m}, \frac{(s^{(\mathrm{m})} + r^{(\mathrm{m})}\Gamma_m)}{\Gamma_m + \theta_m}\right\} \tag{14}$$

where $\Gamma_m = (s^{(\mathrm{b})} - s^{(\mathrm{m})})/(r^{(\mathrm{m})} - r^{(\mathrm{s})})$.

5 Performance Evaluation

In this section, we examine the performance of the proposed routing selection algorithm via simulation. We consider a SDN scenario with the size being 100 × 100 consisting of a SS, a DS and multiple intermediate switches. We assume that all the switches are randomly located in a area. In the simulation, the numbers of intermediate switches are chosen from 6 to 12, the parameters of the arrival curve of the user flow and those of switches are randomly chosen within certain range, as shown in Table 1. For a randomly generated position distribution of switches and the characteristics of arrival curve and service curve, the proposed routing selection algorithm as well as the one proposed in [6] are conducted. The simulation results are averaged over 1000 independent adaptation processes where each adaptation process involves different positions of switches and various characteristics of arrival curve and service curve.

Table 1. Simulation parameters

Parameter name	Value
Peak rate (r^{m}) (Mbps)	1.5–10
Maximum packet size (s^{m}) (Kbit)	1–400
Average rate (r^{s}) (Mbps)	0.7–3.5
Burst tolerance (s^{b}) (Kbit)	38–140
Equivalent service rate ($R_{m,j}$) (Mbps)	10–100
Equivalent latency ($\theta_{m,j}$) (ms)	10–60
The service rate of switches (Mbps)	10–80
The simulation iteration	1000

The simulation results are averaged over 1000 independent adaptation processes where each adaptation process involves different positions of switches. The detailed parameters used in the simulation are shown in Table 1. Fig. 2 shows the effective bandwidth versus the number of switches. It can be seen from the figure that the effective bandwidth increases with the increase of the number of

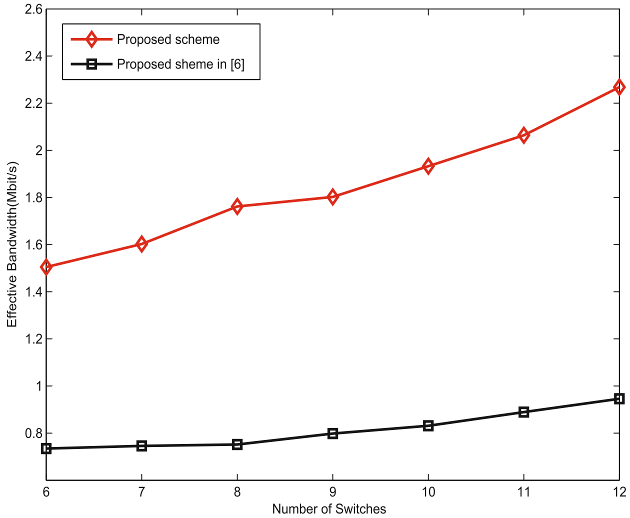

Fig. 2. Effective bandwidth versus the number of the switches

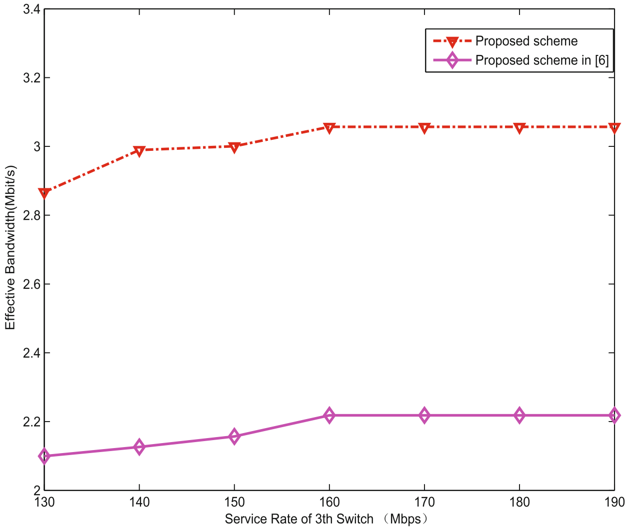

Fig. 3. Effective bandwidth versus service rate

switches, this is because for a large number of switches may offer more options in route selection compared to small number of switches, resulting in better transmission performance in turn. Comparing the results obtained from our proposed scheme and the one proposed in [6], we can see our proposed scheme offers better effective bandwidth than previous scheme. The reason is that our proposed scheme aims of selecting the route with the maximum effective bandwidth, which jointly considers the characteristics of user flow and service capability of switches, including both transmission rate and service delay, while the algorithm proposed in [6] only stresses the transmission rate of switches. In Figs. 3 and 4, we examine the characteristics of one switch on route selection, for instance, we change the service rate and service delay of the third switch, and examine the effective bandwidth of the selected route. In Fig. 3, we plot the effective bandwidth versus the service rate of the third switch. It can be seen from the figure that the effective bandwidth increases slightly with the increase of the service rate of the switch, this is because higher service rate of the switch will offer better transmission performance. Comparing the results obtained from our proposed scheme and the one proposed in [6], we can see our proposed scheme offers better performance than previous scheme.

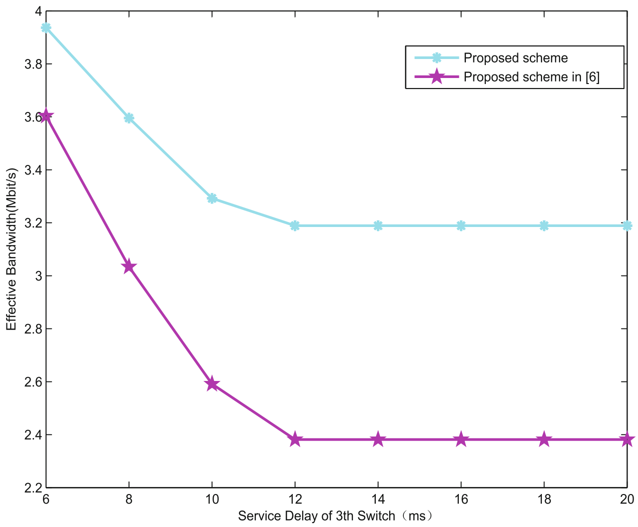

Fig. 4. Effective bandwidth versus service delay

In Fig. 4, we plot the effective bandwidth versus the service delay of the third switch. It can be seen from the figure that the effective bandwidth decreases with the increase of the service delay of the switch for small service delay, then the effective bandwidth will not change clearly with the increase of the service delay. This is because higher service delay of the switch will result in small effective bandwidth. However, as the service delay is relatively high, the optimal route obtained based on the proposed route selection algorithm may not involve the third switch due to its deteriorated performance, thus the achieved effective bandwidth will not change. Comparing the results obtained from our proposed scheme and the one proposed in [6], we can see our proposed scheme offers better performance.

6 Conclusion

In this paper, we propose an optimal route selection algorithm for SDN which aims of achieving end-to-end performance optimization of user flow. Through applying Network Calculus theory, we jointly take into account the characteristics of arrival curve and switches of the network, formulate the transmission performance of user flow as effective bandwidth, and then the route offering the maximum effective bandwidth is selected as the optimal route. The numerical results demonstrate that compared to previous algorithm, our proposed algorithm achieves much better performance in terms of effective bandwidth.

Acknowledgement. This work is supported by the National Science and Technology Specific Project of China (2016ZX03001010-004) and the 863 project (2014AA01A701), the special fund of Chongqing key laboratory (CSTC) and the project of Chongqing Municipal Education Commission (Kjzh11206).

References

1. Matias, J., Garay, J., Toledo, N., et al.: Toward an SDN-enabled NFV architecture. IEEE Commun. Mag. **53**(4), 187–193 (2015)
2. Gantz, J., Reinsel, D.: Extracting value from chaos, Idcemc2 Report, pp. 1–12. Framingham, MA, USA, White Paper, June 2011
3. Adami, D., Antichi, G., Garroppo, R.G., et al.: Towards an SDN network control application for differentiated traffic routing. In: IEEE International Conference on Communications (ICC) (2015)
4. Porxas, A.X., Lin S.C., Luo, M.: QoS-aware virtualization-enabled routing in software-defined networks. In: IEEE International Conference on Communications (ICC), pp. 5771–5776, June 2015
5. Zhang, S.Q., Zhang, Q., Bannazadeh, H., et al.: Routing algorithms for network function virtualization enabled multicast topology on SDN. IEEE Trans. Netw. Serv. Manag. **12**(4), 580–594 (2015)
6. Al-Jawad, A., Trestian, R., Shah, P., et al.: BaProbSDN: A probabilistic-based QoS routing mechanism for software defined networks. In: IEEE Network Softwarization (NetSoft), London, pp. 1–5, June 2015
7. Lee, D., Hong, P., Li, J.: RPA-RA: a resource preference aware routing algorithm in software defined network. In: IEEE Global Communications Conference (GLOBECOM), San Diego, pp. 1–6 (2015)
8. Huang, H., Guo, S., Li, P., et al.: Joint optimization of rule placement and traffic engineering for QoS provisioning in software defined network. IEEE Trans. Comput. **64**(12), 3488–3499 (2015)
9. Le Boudec, J.-Y., Thiran, P. (eds.): Network Calculus: A Theory of Deterministic Queuing Systems for the Internet. LNCS, vol. 2050. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45318-0
10. Shenker, C.P.S., Guerin, R.: Specification of guaranteed quality of service, RFC2212 (1997)
11. Liu, Y., Jiang, Y.: Stochastic Network Calculus. Springer, London (2008). https://doi.org/10.1007/978-1-84800-127-5