SDN 에서 데이터 평면 장애를 해결하는 빠른 우회 기법

Pankaj Thorat, 염상길 추현승
성균관대학교 정보통신대학
Email: {pankaj, sanggil12, choo}@skku.edu

Faster Detouring for Data Plane Failures in Software Defined Networks

Pankaj Thorat, Sanggil Yeom, and Hyunseung Choo
College of Information and Communication Engineering, Sungkyunkwan University

Abstract

Successful deployment of the Software Defined Network (SDN) depends on its ability to cope up with network failures. There are various types of failures that may occur in an SDN. The most common are switch and link failures. It is necessary to recover the network from failures for a continuous service availability. But for the real-time services fast recovery from the failure is required to minimize the service disruption time. In the proposed work, we focused on minimizing the recovery time after the failure is detected. Once the failure is detected, the controller involvement is needed to dynamically reroute the failure disrupted flows from the failed component to an alternate path. The aim of the proposed scheme is to provide a traffic management scheme which can react to the dynamic network events by rapidly modifying the forwarding behavior of the switches for faster in-band network adaptability. The proposed scheme (1) Considers the shared data and control path delay (2) Optimally utilize the network resources (3) Eliminates the need of constant monitoring overhead at the controller which results into faster detouring and ultimately rapid recovery.

1. Introduction

The core idea of OpenFlow is to provide direct programming of a switch to monitor and modify the forwarding nature of the packets. The exchange of control messages between the controller and the forwarding devices happens via In-band channel or Out-of-band channel. In the case of an in-band OpenFlow network, the control traffic (traffic to or from the controllers) is sent on the same channel used to transport data traffic as shown in the Fig. 1 (A). Whereas, in the case of an out-of-band network, the control traffic is sent on a dedicated channel between the controller and switch as shown in the Fig. 1 (B).

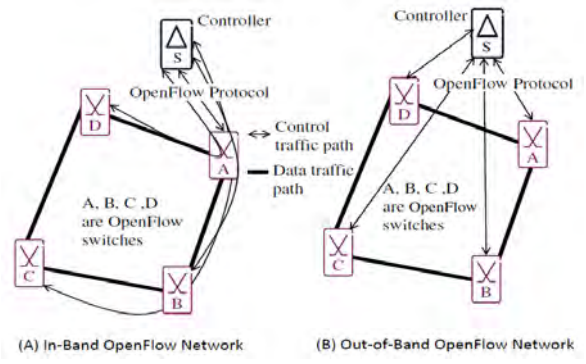

Fig.1: Types of OpenFlow network

Shortcomings of the out-of-band network is that it is expensive to build due to the requirement of an extra physical port on each switch. While in case of In-band control, on the other hand, has the following advantages:

- No dedicated port: There is no need to dedicate a physical switch port to control, which is important on switches that have few ports (e.g. wireless routers, low-end embedded platforms).
- No dedicated network: There is no need to build and maintain a separate control network. This is important in many environments because it reduces proliferation of switches and wiring

For an in-band networks, the control traffic is sent along the same channel as data traffic. The consequences of the in-band channel are:

- Reliability: Excessive switch traffic volume interfere with the control traffic
- Time-sensitive network application: Control traffic generated by the must reach to the destination switch for quick action E.g. the traffic engineering applications like
    - Failure recovery: The controller must find an alternate path and install the flow rules in the alternate path switches to recover the disrupted flows.
    - Load balancing: The controller must distribute the congested flows to optimize the use of network resources
- A rapid traffic management of the congested/disrupted flows must be facilitated to increase the effectiveness of the traffic engineering applications

The rest of the paper is organized as follows: in

Section II, we explained about the effects of the switch/link failure on the network. In Section III, the background work is mentioned. The challenges in the existing scheme is mentioned in Section IV. The proposed idea is presented in Section V. Our plan for experiment is discussed in Section VI. Finally, we conclude our proposal in Section VII.

2. Effect of Node/link Failure on the Network

The most common failures in the present network is either the link failure or the hardware component failure. Therefore for recovery, the system should be sure whether it is a link failure and node failure before performing node failure recovery and the node failure. A physical link in the network fails due to various reasons such as natural disasters, human errors, systems malfunction and equipment problems. A link failure could break the ongoing connection and degrade the network state. Data packets cannot be sent to the failed link. To handle link failure, link recovery is required in which we reroute the failed connections to an alternate path.

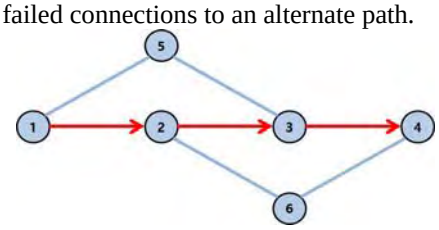

Fig. 2: Before Link Failure

Figures 2, 3 and 4 illustrates how link recovery works. When the link < 2; 3 > of path < 1; 2; 3; 4 > fails the connection between node 1 and 4 disrupts. To reestablish the connection, it is necessary to find an alternate path and reroute the flows to it. If link < 2; 3 > fails the connection can be restored by taking the path < 1; 5; 3; 4 >.

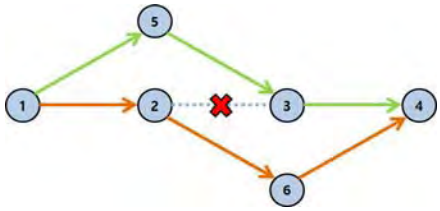

Fig. 3: Rerouting After Link Failure

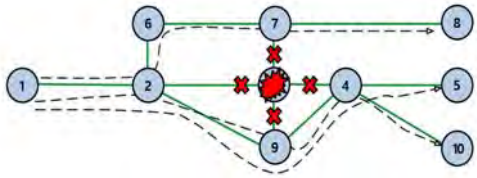

Fig. 4: Rerouting After Link Failure

On the other hand, node failure fails all the links attached to it. When a node fails all the attached links becomes unavailable for sending data. To handle a node failure, node recovery is required in which we reroute the failed connections to an alternate path which is disjoint to the affected links. For example, in figure 3, if node 3 fails, all the links attached to node 3 fails and become unavailable. In that we have to reroute the connection which are on link < 2; 3 >;< 3; 7 >;< 3; 4 >;< 3; 9 >.

Node/link failure affects network and results into degradation in QoS, service unavailability and routing instabilities. Failed node and link can't run correctly and other node still send packets to failed node or link until a new path is found and traffic is rerouted to new path. The continuos traffic inspite of failure affects QoS. Some times when failure occurred, there is no backup path or other path. Host that doesn't have a path to server can't serve service. Routing instabilities is occurred by change in network reachability and topology information. And the need of a new alternate path for maintaining the service.

3. Background Work

There are two background or related works to this topic. We got inspiration from these works so knowing these is helpful to understand our topic. The first one is a 'K shortest paths algorithm'. This algorithm finds K shortest paths. Although it does not gives a perfect result, given paths are eligible to shortest paths. Finding a real K shortest paths requires high cost because shortest path algorithm has to be executed K times. But cost of this algorithm is acceptable because shortest paths are found from intermediate nodes to destination, not from the source. The detailed algorithm is inspired from work mentioned by Marttins et. al.[1].

The other one is a 'Disjoint paths algorithm'. This algorithm finds an edge-disjoint paths. Edge-disjoint path has an advantage of that two paths do not meet. This means that any link failure cannot affect to two paths at the same time. Our proposed algorithm adopted novel method of this algorithm and changed to create a node-disjoint edges to tolerant a node failure.

K shortest paths and finding disjoint paths recover a link failure to use backup paths. In case of single link failure, connection can be rerouted for fast recovery. However, such recovering methods could not support multi-link failures and node failures. Multi-link failures cannot be recovered efficiently by these method. Two different algorithm, enhanced shared-path protection (ESPP) algorithm [2] and self-organizing shared path protection (SSPP) algorithm [3], have been proposed to recover multi-link failures. In node failures, if the node fails, all edges attached to failed node become unavailable. A node failure results into multi-link failure or a disjoint node. Node failure can affect the network performance. Thus, node recovery is required and we provide the recovery to handle.

The current approaches focused on finding the shortest path in the network to reach the destination. But as the network traffic is dynamic in nature, the traffic in the network affects the end to end delay. The hop count based shortest path is not sufficient condition to determine the best path between the switch and the controller for sending the flow modification messages.

4. Proposed Approach for Rapid Detouring

The aim of the proposed scheme is to provide a traffic management scheme which can react to the dynamic network events by rapidly modifying the forwarding behavior of the switches for faster network adaptability. The core goal is to rapidly install/modify the flow rules in the target switches according to the changing network conditions. The proposed

scheme: Considers the shared data and control path delay; optimally utilize the network resources; Eliminates the need of constant monitoring overhead at the controller which results into faster detouring and ultimately rapid recovery

The research problem are as follows: (1) Congested control path: Solution: Use multiple control paths to distribute the flow modification control messages to the switch (2) Continues monitoring to determine the lightly congested path: Solution: Alter the paths based on the congestion on the datapaths (3) Dynamically respond to the changing network statistics: Solution: The OpenFlow controller sends a Barrier Request message to request that the OpenFlow-enabled switch complete processing of all messages sent before the Barrier Request message before processing any messages sent after the Barrier Request message. This ensures that the virtual switch processes all message dependencies and sends all notifications for completed operations before proceeding with new requests.

The proposed scheme works as follow:

1. Detect the anomaly (Congestion/failure) in the network.
2. Find the detouring plan
   a. Calculate the alternate paths for the flows.
   b. Find the target switches (T_{switch}) where the flow addition is required.
3. Find the shortest possible control paths between the controller (C) and the T_{switch}
4. At time instant t=0,
   a. Transmit 1000 flow addition messages on each control paths cp_i
   b. Transmit the 3 barrier request message to T_{switch} on each cp_i and wait for the barrier_reply message from the T_{switch}
5. After receiving the barrier_reply message from the T_{switch}
   a. Calculate the difference between the time at which the first flow addition message was sent and the barrier_reply message received (reroute_delay)
   b. The control path with the smallest reroute_delay is considered as the less congested and more control traffic is sent on it compared to the remaining control paths

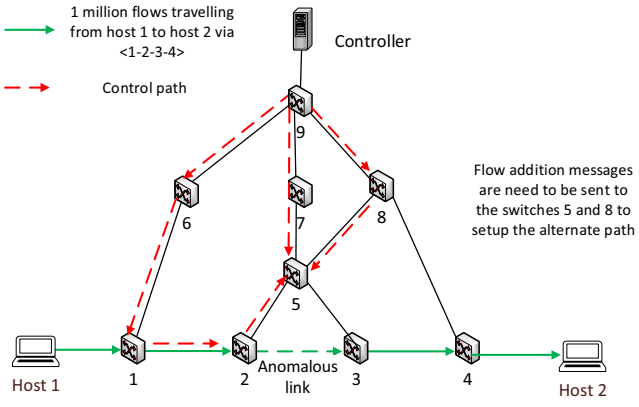

Fig. 5: Optimal datapaths for faster control. The controller decides the best possible control paths between the controller (C) and the T_{switch} (At least 2, maximum 3): cp1: <C-9-6-1-2-5>, cp2: <C-9-7-5>, cp3: <C-9-8-5>

5. Plan for the Evaluation

In this section, we are discussing about the evaluation technique of the proposed scheme. This part is Design of the Experiment. First, we should be created a graph which is based on ATT topology [4]. It is consist of 25 nodes, 52 links. Therefore, we assume that the size is enough for verification of the proposed scheme. The topology is shown below in figure 6.

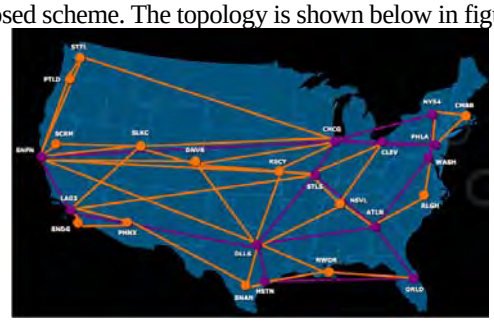

Fig. 6: ATT topology

For evaluation, we are planning to create ATT topology and create connections from certain set of source and destination. By simulation we will fail a switch or a link and reroute the connection which are affected by the node/link failure from primary path to the backup path. For faster detouring we will use the best control paths using the proposed scheme to rapidly deliver the control traffic. We will compare the proposed scheme with equal cost multiple path algorithm. We will present the analytical model for the proposed scheme.

6. Conclusion and Future Work

The advantage of the proposed scheme is that it minimizes the monitoring overhead for finding least congested path. It reduces control traffic in the network and sonsiders the shared data and control path delay. The proposed schme optimally utilize the network resources and suitable for time-constrained applications in an in-band network like load balancing, failure recovery. The proposed scheme works for only in-band OpenFlow network.

Acknowledgement

This research was supported in part by PRCP (NRF-2010-0020210), ICT R&D program (B0101-15-1366, Development of Core Technology for Autonomous Network Control and Management), and G-ITRC support program (IITP-2015-R6812-15-0001), respectively.

References

[1] Martins, E.D.Q.V., Santos, J.L.E.D.,"A new shortest paths ranking algorithm". In the Investigacao Operacional 20(1). 2000 (pp. 47–62).

[2] Guo, L., X.Wang, J., Cao, X., Zheng, A.,"Recovery escalation with load balancing and backup resources sharing in survivable WDM optical networks". In the Photon. Network Commun. 18. 2009 (pp. 393-399).

[3] Zheng, W., Liu, S., X. Qi. "Multi-links recovery algorithm in survivable WDM optical network". In the Photon. Network Commun. 21. 2011, (pp. 1-6).

[4] ATT topology. https: //www:att:com/Common/merger/files/pdf/wired-network/Domestic0C-768Network.pdf