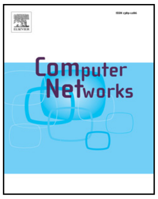
Computer Networks 120 (2017) 56–70

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet



Global Flow Table: A convincing mechanism for security operations in SDN

Xiaofeng Qiu, Kai Zhang*, Qiuzheng Ren

Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing, China

ARTICLE INFO

Article history:
Received 28 October 2016
Revised 13 March 2017
Accepted 3 April 2017
Available online 6 April 2017

Keywords:
Security operations
Software defined networking
Global flow
Weak vertex cover

ABSTRACT

One of the key challenges of network security is that security middle boxes, such as firewalls and Intrusion Detection Systems (IDSs), only have local view of the network. This lowers the efficiency of security detection and makes it difficult to locate the sources of the threats. There have been growing demands for security operations and appliances that are aware of the distribution and behavior of flows in the whole network; logically centralized control ability of Software-Defined Network (SDN) makes it possible for the network controller to acquire the global view of the network. In this paper, we propose a mechanism named Global Flow Table (GFT) which can provide security appliances and operators with paths of all the flows in SDN network, in addition to their sources, destinations, setup and terminate time, traffic volume and directions. A weak vertex cover based GFT algorithm which sacrifices less than 5% accuracy is also provided to improve scalability. Tests with different network topologies of cloud computing center and enterprise networks show promising performance. Utilizing the Global Flow Table, we built several applications to illustrate how GFT could benefit the security operations.

1. Introduction

Recently, profound interest has been developed in automating network control, particularly with the emergence of software-defined networks (SDN). SDN is a new software-based network architecture and technology. Loose coupling between control plane and data plane, centralized control of network state and transparency of underlying network facilities for upper-layer applications are the most important features of SDN. Network intelligence and state are logically centralized by network controller, which provides a global view of the whole network [1,2]. SDN switches are centrally managed through a well-defined southbound interface such as OpenFlow [3] which allows control plane to write forwarding rules to switches. OpenFlow is the only certified southbound interface by ONF (Open Networking Foundation), which has a pivotal position in the development of SDN.

Similar to the design philosophy of SDN, network security operation is also on its way to automation and intelligence. Besides, it also requires self-healing and real-time anomaly detection [4]. However, unlike their counterparts in network operation, security operators are still lack of tools or applications to obtain the global view of network behaviors.

In this paper we propose a new mechanism—Global Flow Table (GFT). It generates and stores the complete path of each flow containing the direction, passing switches, forwarding ports and traffic volume in the current network. It can make up the deficiency in security detection. For example:

(1) With the assistance of GFT, path abnormal attacks in network can be easily detected with simpler and faster algorithms. Although detection based only on GFT will provide relatively high false positive rate for some kinds of attacks, the GFT will greatly help to reduce the burden of security detection as only the suspicious flows or nodes will be reported to more sophisticated security devices for further fine grained detection. In Section 6.1, we demonstrate how suspicious Man-in-the -middle attack (MITM) that intercepts and inserts into the flow path can be intuitively detected utilizing the GFT. Besides, the source address forgery attack could also be detected by comparing the global flow pathes in GFT with information of links obtained by the northbound API of the SDN controller. In addition, Denial of Service (DoS) attackers [5] that drop packets maliciously in data plane of SDN can also be coarsely detected by checking the forwarding port item of the GFT.
(2) GFT could help to improve the efficiency of security protection against Distributed Denial of Service (DDoS) attack. By learning the passing path of each network flow provided by

* Corresponding author.
E-mail address: kaibao@bupt.edu.cn (K. Zhang).

http://dx.doi.org/10.1016/j.comnet.2017.04.002
1389-1286/© 2017 Elsevier B.V. All rights reserved.

GFT, the Intrusion Detection Systems(IDS) can not only handle the flow in detection point, they could also notify the switches closer to the malicious sources to drop malicious flows in order to suppress attacking traffic.

(3) GFT could help to detect attacking flows with low traffic failed to trigger the threshold of IDS. For example, slow scanner scans tenant networks in a round-robin fashion and DDoS attackers launch attacks targeted to the victim from lots of Bots in order to keep the IDS not triggered. GFT will help the security appliances to detect these kinds of attackers that elaborately designed to defeat the threshold based detection algorithm. In Section 6, we demonstrate two applications upon GFT, the Global Flow Monitor and Global Flow Path visualization, both provide security operation with roughly suspicious flows or nodes that are eligible for further sophisticated detection.

GFT not only helps security operations, but also makes network operation and management more intelligent and efficient. Commercial network troubleshooting tools provide visibility through packet sampling [6,7], configurable packet duplication [8–10], or log analysis [11], etc. Most of the work mentioned above lack network-wide visibility and packet-level state consistency which are provided by GFT.

GFT is built based on flow tables of the SDN network. In the SDN architecture, control and data planes are decoupled. The network controller chooses optimal network path for application traffic through issuing flow table entries to each switch. All packets processed by a switch are compared against the flow table entries in this switch. Flow table entry contains a set of matching entries, activity counters, and a set of zero or more actions that the switch will conduct on the matching packets. The matching entry acts as identity of a flow in a flow table. Through identifying flow table entries, we can find the specific flow passing different switches along its path [12]. Therefore, flow table entries existing in SDN networks form the basis for building GFT.

The key challenge in creating GFT lies in flow information collection and path computation of all the flows in the whole network. The first step of creating GFT is to gather flow table entries. Flow table entries in switches could be read by the SDN controller through OpenFlow protocol between the controller and switches. GFT could further get these flow table entries through controller's northbound API. As introduced previously, during the flow setup, the controller could record most items of a flow table entry before issuing it to the switch. However, in order to acquire complete visibility in a large SDN network, the controller has to collect real-time statistic information in counters of flow table entries at each switch. Such fine-grained control and visibility come with two kinds of costs: switch-implementation cost involving the switch's control-plane and distributed-system cost involving the controller. As a result, collecting flow information from all the switches via the pull-based Read-State mechanism can actually create dreadful control-plane load. Meanwhile, GFT need timely access to statistics to accommodate the dynamic network. Moreover, statistics gathering competes for limited bandwidth between switch and controller for flow setup. The more frequently statistics are gathered, the fewer flows a switch can set up. Hence it is necessary to reduce the cost of collecting flow table entries while building the GFT [13].

In this paper, we proposed a weak vertex cover based GFT algorithm to find the optimum Collecting Set of switch nodes instead of gathering all the flow table entries from every switch using the Vertex Cover (VC) problem with flow-conservation law [14] to collect flow table entries. A VC is a set of nodes in a graph that every edge of the graph has at least one endpoint in the set. A minimum cover is the VC which has the smallest number of nodes for a given graph. According to the traffic of each edge associated with the VC set, we can determine the flow of arbitrary edge in the network. Since VC is a NP-hard problem, accordingly, it is required to reduce the number of effective measured set in order to find the minimum effective measuring set. Using a VC of the network graph to determine the set of nodes on which the controller collects the flow tables can obviously result in a substantial reduction in the number of measured nodes. Moreover, it is possible to be more effective by exploiting the flow-conservation law. Based on the flow table entries gathered from the Collecting Set of network nodes, all the flows are visible and the path of each flow can be computed. Further, it greatly reduce the cost of statistics gathering.

Path computation is also a key factor in the GFT. Flow table entries collected are only a set of flows, which cannot reflect the path information and traffic distribution. In this paper, we design a global path calculation algorithm to compute the complete path of all the flows existing in the current network. GFT exposes a restful API for applications to specify, receive, and act upon global flow information of interest. Based on the GFT, we developed several applications such as the Global Flow Graph, Global Flow Path visualization, MITM Detection application and Global Flow Monitor, which can significantly promote efficiency and performance of security operations. Details of these applications will be introduced in Section 6.

The contributions of this paper are as follows:

1. We propose a new mechanism named GFT and corresponding algorithms to build the global information of the entire network flows, which makes the security appliances aware of network behaviors and the security operations intelligent.
2. To reduce the cost of gathering the real-time information of flows, we harness the Weak Vertex Cover problem to find the optimum Collecting Set of measured switch nodes. Statistics of all the flow are only pulled by the controller from the switches in the collection set which greatly cut the cost of GFT.
3. Based on the collected flow table entries, we propose a global path calculation algorithm to recover and compute the path of each network flow. We also design a data structure to store GFT for the purpose of searching information on specific flows conveniently.
4. We developed several applications based on GFT, such as Global Flow Graph, Global Flow Path visualization MITM Detection application and Global Flow Monitor. These applications help network security managers to detect and trace the source network attacks. Meanwhile, they can also guarantee the correctness of the network behaviors.
5. In weak vertex based GFT, scalability is at the cost of accuracy, to apply GFT to different topologies of cloud computing center, campus and enterprise network, we carried out experiments in the scale of these networks. In consequence, the accuracy ratio of GFT is always more than 95% with different network topologies and scales when adding marginal nodes. The time delay is less than 1000 ms when there are 300 switch nodes in the network and each of the nodes contains 100 flow table entries.

The remainder of this paper is structured as follows: Related work is presented in Section 2. Global Flow Table is introduced in Section 3. Section 4 presents weak vertex cover based GFT algorithm. Section 5 is the experimental results and evaluation. The applications developed upon GFT are described in Section 6 and we conclude the paper in Section 7.