Multi Controller Software Defined Network Link Fault Location Based on Tree Decomposition Method

Xing Ye[1(✉)], Xing-Guo Luo[1], Yu-Xiang Hu[1], and Min Lin[2]

[1] China National Digital Switching System Engineering and Technological R&D Center,
Zhengzhou, China
Yxeingyx358@163.com
[2] Henan Institute of Animal Husbandry Economy, Zhengzhou, China

Abstract. In order to overcome the problem that the existing scheme is difficult to adapt to the network link additions and deletions to the situation, this paper proposes a link fault localization scheme based on tree decomposition. Firstly, a multi degree lossless image segmentation algorithm is proposed to divide the network into several independent sub networks. Secondly, the maximum width priority tree decomposition algorithm is proposed to generate multi loop in the sub graph. The average number of links and the total delay of the proposed algorithm are better than that of the genetic algorithm. Experiments show that the number of indicators such as static flows in the schema needed to be sent to the table and detection information needed to be sent to detect link faults is superior than the existing programs.

Keywords: SDN · Link fault locator · Decomposition tree · Image segmentation

1 Introduction

Software-defined network is a new network architecture model. It will separate control and forwarding network, through the open software defined API to realize the flexible reconfiguration of the network function [1]. With the use of the SDN in the WAN, the control layer using a single controller cannot complete the statewide network monitoring and management. To achieve scale control layer, a plurality of controllers together to complete the whole network management and monitoring become the development direction of large-scale deployment [2]. Although the multi controller can achieve the management and control of large scale software defined network, it also brings some new problems. For example, in a multi controller software defined network, each controller only manages a portion of the network. Multiple controllers need to exchange the network topology information in their own control domain. In this way, the whole network topology information can be generated by each controller. On the basis of the application of the controller, a variety of resource management and allocation decisions are made. But when a controller and its control domain switches between communication fault, or a controller of its control domain switch control ability is weak, the controller will not be able to correctly, efficiently obtained body control network

topology information in the domain. Therefore, it is very important to have a full network topology information for a controller.

2   Research Status

In a multiple controller, how to achieve a single controller link failure detection and location in the network and then get the complete information of the entire network topology is a new topic.

Ulas proposed a single controller to use its own master network topology information to generate all links in the single loop overlay network [3]. And in order to send the OpenFlow switch to the static flow table entry (the number of flows close to the number of optimizations), the single loop will be generated in the OpenFlow switch network. The controller can detect the topological state of the network by transmitting the detection information to the single ring, and it is able to detect and locate the link fault in the network by forwarding speed of the data path. At the focus of the study is to ensure the generation of single-ring overlay network on the basis of all the links, so that repeated single ring path contains the minimum number, thereby reducing the number of controllers to switch delivers static flow rules. On this basis, the author analyzes the number of single-loop positioning network link failure.

Khurshid main research is how to monitor the real-time flow rules issued, modification and deletion influence to the whole network status [4]. The authors suggest adding a "VerilFlow" layer between the control and data forwarding layer, which is the first function of an existing network into multiple equal subnets, and access to each subnet data forwarding topology.

Wundasm describes how to locate network problems in a network switch or router configuration problem caused [5]. This paper present OFRewind modules, it can achieve large-scale, centralized control of the network and the continuity of the recording and replay functions based on the SDN network. Through these two functions to identify the control software failure, data forwarding in the link bottleneck and positioning switch or router configuration fault.

Hongqing Zeng and Changcheng Huang heuristic spanning tree algorithm to generate a network topology tree, it not included in the tree and the tree link connection generates a detector ring [6]. Based on the generated ring signal generator to be placed and received signals, signal generator sends a detection signal to the ring to locate the all-optical network link failures.

Ying Xuan mainly to solve how to locate Optical Networks multi-link failure problems, this paper proposes a centralized tree decomposition algorithm and random walk localization algorithm [7]. Centralized decomposition algorithm suitable for small networks, the algorithm generates network topology tree subgraph, and the signal generator is placed on the root node of the tree, and place the signal receiver on the leaf nodes of the tree. The Random walk algorithm is suitable for large networks, the algorithm randomly generated multiple paths in large networks. Until the resulting path contains all the links in the network, the algorithm terminates. Although the two algorithms can

locate the multi-link failures in the network. But the number of the signal generator and the signal receiver need to be installed, and the economic cost is large.

In the process of summarizing and analyzing the above papers, it is found that the cost standard of link failure in all optical networks is not the same as measuring the link failure in the network. In all-optical networks is to detect signal generator and receiver of a number, and the size of the detection channel occupancy of the light spectrum as the standard. The measure in the SDN network is the number of static flow rules established on the switch, and the number of detecting control information.

3 Methodology

3.1 SDN

SDN is a new kind of network technology by separation of control and data network equipment, network abstraction of the application program interface provides to the application layer, so as to construct the open network environment programming [8]. On the basis of various underlying network resource virtualization, to achieve the
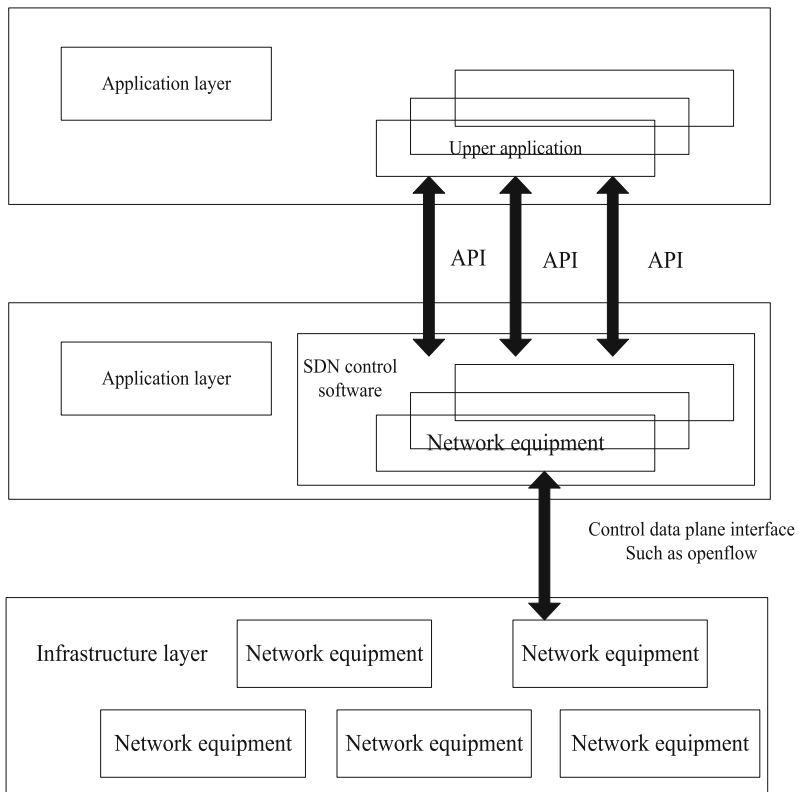

Fig. 1. Architecture of SDN

centralized control and management of the network. Compared with the traditional network equipment, the control ability of network equipment is centralized to the controller. Through the network operating system software-driven approach to achieve a flexible, highly automated network control and service configuration. The logical view of the SDN schema is shown in Fig. 1.

3.2 Tree Decomposition

Tree decomposition approach is to figure each node gradually removed from the drawing, and edge connected thereto, together to save the stack [9]. The rule is deleted from the beginning of the implementation of the lowest node delete. After deleting a node, it need to verify that the two sides connected between the node's neighbors, if not then build such edges.

3.3 Multi Link Fault in Multi Controller Based on Tree Decomposition

3.3.1 The Basic Idea of Link Fault Location Scheme

Firstly, the network topology graph is decomposed into a set of multiple sub graphs based on the tree decomposition of multi degree lossless graph partitioning algorithm. It can be abstracted into a tree structure in Fig. 2(a). In graph (a), a SDN topology graph G is decomposed into a number of interconnected sub networks $Gi, Gr. Gn. Gi$ represents the i sub network in figure G. G in the topology diagram with n controller, wherein the i-th control domain controllers' Ci own one or more of OpenFlow switches in the subnet Gr (switch control domain can be distributed simultaneously in the different sub net), the controller Ci in order to detect the link status of the whole network, it needs to send a probe message to other subnets.
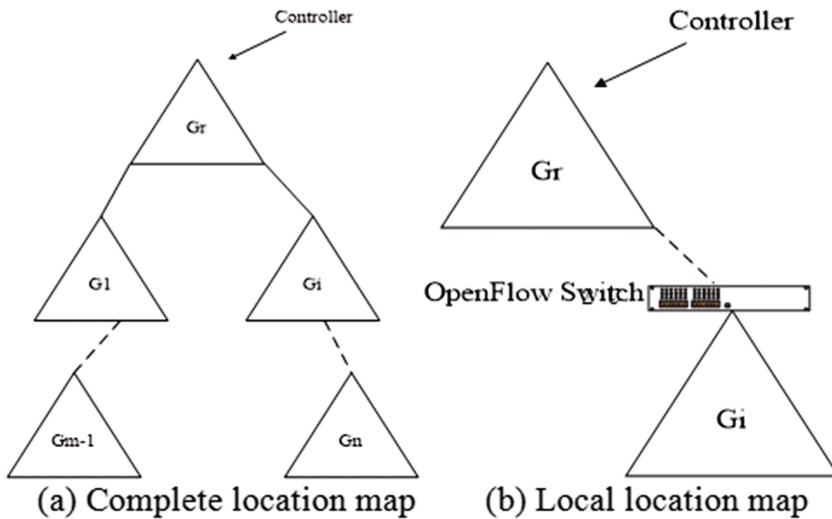

Fig. 2. Scheme of locating link failure

There are two ways of communication between controller and sub network. The first is: If a subnet Gi contains the controller in the control area of the switch, the controller chooses the switch as a proxy switch that represents the controller through the switch and the sub network communication. The controller can send the detection information to the sub network Gi through the proxy switch. Another way is: If a sub network Gi does not contain the switch under the controller Ci control domain, Ci through the connectivity between the sub network to find a way to reach the shortest path Gi border switches. The path to create a link controller Ci and Gi switch between subnet boundaries, the boundaries of the switch as a proxy switch, the controller sends a probe message to the subnet through the proxy switch. It is shown in Fig. 2(b).

3.3.2 Multi Degree Lossless Image Segmentation Algorithm Based on Tree Decomposition

If the direct method in accordance with the decomposition tree topology graph into a tree topology, each tree node as a subnet have the following problems: Firstly, through the tree decomposition, the number of nodes is too much, each tree node contains a fewer vertex tree, the independence of the tree node is weak, connectivity is too strong. Second, there is no link between the vertices of some tree nodes. In order to solve the problems existing in the sub network generated by the traditional tree decomposition method, a new method of multi degree lossless segmentation is proposed. The algorithm is based on the traditional tree decomposition method adds the tree node merging step.

Abundance lossless decomposition algorithm details are as follows: First, the traditional tree decomposition algorithm is the exploded tree, tree node contains the original set of vertices and edges. Then the tree merging algorithm is used to combine the tree nodes to meet the defined sub graphs.

The specific idea of tree node merging in multi degree lossless decomposition algorithm is as follows: according to the traditional tree decomposition method, a topological graph is generated. Starting from the root node of the tree, using a breadth first search algorithm each node's child nodes. When the child nodes contain the vertices and edges meet certain conditions, nodes of the child nodes and above can do not rely on the idea tree nodes. It can be merged into a meet the definition of a new node. The node and its subtree nodes can also does not depend on the child nodes of the master node and the node above can constitute a new node to meet the definition, which we call the child of the node is the critical nodes. For critical nodes in the subtree of the root node, the same with the above ideas to merge, until incorporated into the tree leaf nodes.

3.3.3 The Maximum Degree-Breadth-First Spanning Tree Algorithm (MBIST) Construction of Polycyclic

For topology G, initialize an empty tree T. To select the figure G to moderate the maximum vertex as the root node r of the tree T. Use the width priority principle to connect the neighbor vertices of the graph with the root node r as the node of the root node to be added to the tree T, and the degree of the node in the tree is updated. Then, depending on the size of the tree T contains a node, select the node degree, its neighbor in the original vertices as its child node is added to the tree (already included in the tree

T neighbor nodes do not need added again). When the neighbor vertices of all the nodes in the same depth of the tree *T* are added to the tree *T*, the degree of the node in the tree *T* is re updated (excluded when calculating the degree of a node has been included in the side of the tree). Finally, repeat the third step operation, until all the vertices in the original image are included in the *T* tree, the end of the algorithm.

3.3.4   Algorithm Performance Evaluation

The performance of the maximum depth first algorithm and other algorithms are compared, and the comparison of the index number of links is balanced and the number of links contains the total number of links. The FST and GST, BFS algorithm was compared. It is shown in the Table 1. For the NSFNET network, MBFST is better than HST and BFST in the uniformity of the ring, it's mainly in the difference of the ring contains the maximum number of links and the minimum number is smaller than the other two algorithms. The total number of links contained in the generated all rings is also less than HST and BFST. For BELLCORE network, MBFST and HST generated balance polycyclic links and the total number of the link are same, but higher than BFST performance.

Table 1.   Maximum degree-breadth first multi-cycles, creating an algorithm

Network topology	Algorithm	Single ring contains the maximum number of links	Monocyclic contains the minimum number of links	All total number of links contained in a ring
NSFNET	MBFS	6	5	41
	HST	7	5	47
	BFS	7	3	49
ARPA2	HBFS	13	6	43
	HST	22	6	63
	BFS	15	6	46
BELLCORE	MBFS	7	3	63
	HST	7	3	63
	BFS	9	3	82

3.3.5   Subnet Failed Link Locating Scheme

Link coding.   First, the generation of L-N + 1 ring contains a diagram of all links, and the figure of a link may appear in a number of different ring. Let e be a graph $G = (V, E)$ in a link, figure *G* is the ring set $(C_{i \in 1}, I = 1, \ldots, L-N + 1)$ coverage, which link e may appear in the *k* ring. Secondly, when the controller through the border switch or switches the agent to send all of the rings in the detection signal, if only *e* link fails, the ring contains *k* link *e* cannot detect feedback to the agent switch, and does not include *e* ring link can normally detect feedback information. If the proxy switches to get a feedback loop, the loop is marked as 1. If the proxy switch does not get a feedback loop, the loop is marked as 0. Then, we can obtain a combination of values $v = \{v_{C1}, v_{C2}, \ldots, v_{Ci}, \ldots, v_{CL-N+1}\}$, where v_{Ci} represents the state of the *i*-th loop has a value of 0 or 1. After using MBFST

algorithm network topology NSFNET polycyclic generated, according to the encoding scheme to each link series of fault codes shown in Table 2. The table on the left represents the polycyclic state code corresponding to the combination of value, the table on the right represents the polycyclic get left state code corresponding to the combination corresponding to the value in the network of the failed link.

Table 2. Coding effect

Information code								Failed link
C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	
0	0	0	0	0	0	0	1	9–13
0	0	0	0	0	0	1	0	10–14
0	0	0	0	0	1	0	0	9–14
0	0	0	0	1	0	0	0	10–13
0	0	0	1	0	0	0	0	1–4
0	0	1	0	0	0	0	0	8–9
0	1	0	0	0	0	0	0	8–7, 4–7
1	0	0	0	0	0	0	0	1–2
1	1	1	1	0	0	0	0	3–6
0	0	0	1	1	0	1	0	4–6
0	0	1	0	0	1	0	1	6–11, 9–11
0	0	0	0	0	1	1	1	6–12
1	0	0	1	0	0	0	0	1–3
0	1	1	0	0	0	0	0	4–5
0	0	0	0	1	0	0	1	12–13
0	0	0	0	0	1	1	0	12–14
0	1	1	0	0	0	0	0	2–8
0	0	0	0	1	0	1	0	4–10

In the above table, only a single link fault code is listed, and the corresponding code word is also listed for multi-link fault. If the code word is still unique, the code word can correspond to a multi-link fault combination.

Acquisition of code value. In order to detect the link state in a subnet, the controller is still sending a probe packet to the proxy switch. But for each subnet, the controller sends only one probe message, not L-N + 1. After receiving the detection information, the proxy switches not only forwards the detection information to a ring in the detection information, but also uses the multicast mode to transmit the detection information to all the loops through it. Use the flow table entry group operating function Openflow switch, when the switch multicast will be sent to a different ring probe packet marked with different ID numbers. When the controller receives the information detected by analyzing the ID number can be returned to the ring this information. A specific process is as follows: when the controller generates the ring to send flow table under static switches, they are set to a static set of actions for each flow table entry. The content of the group is to operate all through the flow table entry packet source *ip* address set to ring ID number. Because OpenFlow switches support group operation is relatively

simple, one of the functions supported is to change the function *ip* packet's source address. In the design of ID, the design of the LAN IP address format. When a ring OpenFlow switch receiving the probe packet through the port number to match the flow table entry, the switch through the flow table entries pre-set group of operations required to forward packet's source *ip* address portion of the ring ID number. It is shown in the Fig. 3.
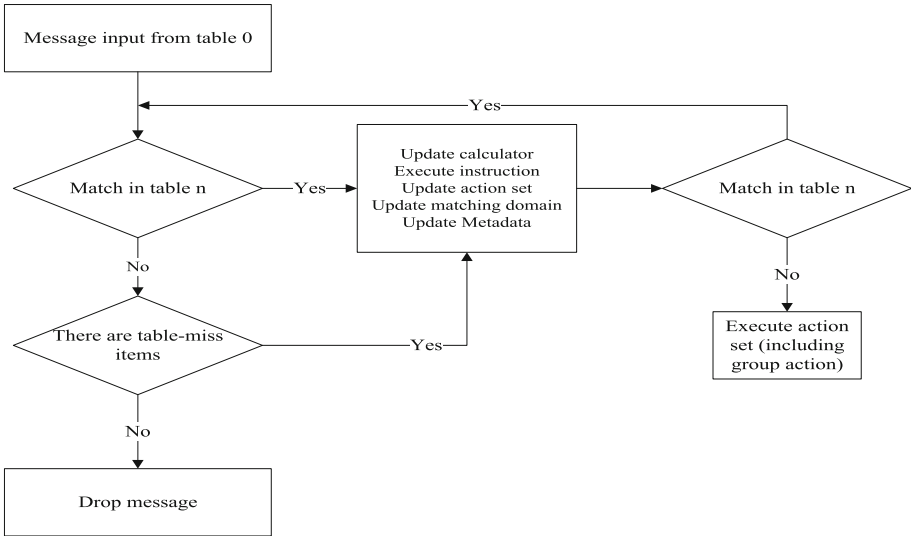

Fig. 3. Procedure of addition of ID number

4 Discussion

Verify that the topology graph is obtained from the topologyzoo web site to obtain the real map, respectively, from the selected SMALLNET, APARENET, BELLCOM, TW_TELECOM four types of network topology. The degree of nodes in SMALLNET is generally greater than or equal to 3. BELLCOM, TW_TELECOM, APARNET in the node degree is generally 2 to 3. The characteristics of BELLCOM and APARNET network topology are different. In that the distribution of node degree, BELLCOM is larger than that of APARNET. TW_TELECOM is larger than the size of BELLCOM and APARNET.

Each tree node (the sub network) is selected as a boundary switch, the controller only needs to send a packet to the border switch. The boundary switch transmits the detection information to the ring which is composed of a static flow table item in advance, and the detection information is carried out by the group operation set in the flow table. Although the probe multicast packet in the process to increase the number, but the whole process is the switch probe information automatically, not sent to the switch controller. Therefore, the tree decomposition method to detect the link fault in the network. It is needed to detect the information tree and the number of border switches,

which is related to the number of the sub network generated by the topology prepro-cessing. The single-loop formation, double loop generation algorithm based on a genetic algorithm generation algorithm, and an algorithm for generating the second ring based on genetic algorithm are to use the dichotomy probe generation ring link failure, so the number of its probe packets, generated by the number of rings and each ring of the size of the impact.

5   Conclusion

The complete network topology information is required for the controller to carry out the functions such as path selection, load balancing, path optimization, traffic distribu-tion and so on. In the multi controller, if there is a problem of information exchange between controllers, a single controller cannot get the correct and complete network topology information [10–12]. The decisions made by each controller, it may conflict with each other, and the paralysis of the entire network is more serious. For resolving this problem, we propose a method based on tree decomposition to locate link faults in a multi controller. In this paper, we propose a method based on tree decomposition for multi degree lossless image segmentation and the maximum breadth first spanning tree. In this algorithm, the network topology is decomposed into several sub networks, and the degree of vertices in the sub graph can be guaranteed to be more than or equal to 2. In the spanning tree topology, it needs to ensure that the depth of the minimum spanning tree at first. Secondly, in the spanning tree, it needs priority to place original moderate large vertex, then according to the tree generated polycyclic. The experimental results show that the link failure in the network based on tree decomposition method is better than other algorithms in detecting the number of packets and the number of static flow table items. But it is more suitable for the single link fault or chain link fault location. It is weak for more than three link fault location ability.

In this paper, we proposed solutions for a single controller independent to ascertain the status of the network topology, and positioning in the whole network link failures. It needs to be further improved. It has a very good adaptability to dynamic network. At the same time, the most generous - width priority spanning tree algorithm generates a number of multi – loop, it can greatly reduce the number of present static flow table items. And it also can reduce the number of detection information for detecting link failures, but it is not strong enough to locate multiple faulty links. It needs for improve-ment in the following studies.

References

1. Feamster, N., Rexford, J., Zegura, E.: The road to SDN: an intellectual history of programmable networks. ACM Sigcomm Comput. Commun. Rev. **44**(2), 87–98 (2014)
2. Cvijetic, N., Tanaka, A., Ji, P., et al.: SDN and OpenFlow for dynamic flex-grid optical access and aggregation networks. J. Light. Technol. **32**(4), 864–870 (2014)
3. Kozat, U.C., Liang, G., Kokten, K.: On diagnosis of forwarding plane via static forwarding rules in Software Defined Networks. Proc. INFOCOM **12**, 1716–1724 (2013)
4. Khurshid, A., Zhou, W., Caesar, M., et al.: Veriflow: verifying network-wide invariants in real time. ACM Sigcomm Comput. Commun. Rev. **42**(4), 467–472 (2012)
5. Zeng, H., Huang, C., Vukovic, A.: A novel fault detection and localization scheme for mesh all-optical networks based on monitoring-cycles. Photonic Netw. Commun. **11**(3), 277–286 (2006)
6. Xuan, Y., et al.: Efficient multi-link failure localization schemes in all-optical networks. IEEE Trans. Commun. **61**(3), 1144–1151 (2013)
7. Li, B., Moataz, F.Z., Nisse, N., et al.: Minimum size tree-decompositions. Electron. Notes Discrete Math. **50**, 21–27 (2015)
8. Yang, M., et al.: Software-defined and virtualized future mobile and wireless networks: a survey. Mob. Netw. Appl. **20**(1), 4–18 (2014)
9. Sahani, D.V., et al.: Magnetic resonance imaging biomarkers in hepatocellular carcinoma: association with response and circulating biomarkers after sunitinib therapy. J. Hematol. Oncol. **6**(1), 51 (2013)
10. Mary, V.G.: A hybrid dynamic routing protocol for congestion free traffic in vanets. J. Inf. Technol. Rev. **7**(2), 52–59 (2016)
11. Rujuta, P., Namrata, V., Ami, V., Khushali, D.: User future request prediction using F.P tree. J. Intell. Comput. **6**(1), 26–36 (2015)
12. Carrera, Ã., Iglesias, C.A.: Exploiting structural knowledge using network description language and causal models for fault diagnosis in wireless sensor networks. J. Netw. Technol. **6**(4), 91–103 (2015)