2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC) | 978-1-7281-5244-8/20/$31.00 ©2020 IEEE | DOI: 10.1109/ITAIC49862.2020.9339189

Research on abnormal traffic diagnosis based on deployment mode of firewall

Ming Ren[1] ,Weiyan Zhang[1] ,Siqi Kong[1] ,Dali Zhou[1] ,Danping Li[1] ,Yanhui Tian[1]

1.Institute of Computer Applications, CAEP, Mianyang, China

417376813@qq.com, 5749369@qq.com

Corresponding Author:Siqi Kong Email:372857202@qq.com

***Abstract*-The firewall based on quintuple for access control plays a very basic and vital role in network security protection, but the diversified firewall deployment mode also brings challenges to the network configuration: unreasonable network configuration may cause that the network traffic is interrupted, and when the interruption occurs, it is impossible to find the cause of the interruption if only analyzing the switching and routing involved in the network configuration. This article first analyzed the forwarding characteristics of traffic passing through the firewall, then proposed four types of firewall traffic forwarding models. No matter how the firewall deployment mode changes, the way of the traffic passes through the firewall must belong to one of the four models. This article further analyzed the four types of forwarding models based on the characteristics of firewalls, and found that three of them can cause abnormal firewall traffic. For this reason, this article proposed an algorithm for diagnosing abnormal models, and provided a solution based on network technology and firewall characteristics.**

Keywords- Firewall deployment mode; firewall traffic forwarding model; Abnormal traffic; Anomaly model detection algorithm;

I. Research background

The current global network environment is becoming more and more complex. Various types of network attacks are increasing year by year and showing a diversified trend. Attack methods are emerging one after another, and protection methods are becoming more and more diverse, including WEB application firewalls, user online behavior analysis,intrusion prevention systems, firewalls, user authentication equipment and so on, among which the firewall plays the most basic protective role in the entire network structure and is often used as the first line of defense for security protection. At present, various researches are mainly focused on firewall packet classification algorithms[1][2], firewall architecture design[3][4][5[6]], firewall strategy[7][8] etc. The research on the impact of firewall deployment mode on network traffic and network settings is still in a blank area, However, after the firewall is deployed online, it has a obvious impact on network traffic and network configuration. If you do not analyze and prevent in advance, it may cause some network traffic to be interrupted after going online, which will have a catastrophic impact on the business. It is in such an environment that this article first proposed four forwarding models of traffic passing through the firewall based on analysis of firewall network traffic, and proposed an algorithm for diagnosing abnormal traffic models, and provided a solution based on network technology.

II. FIREWALL TRAFFIC MODEL

2.1 FIREWALL TRAFFIC MODEL

Since the main function of the firewall is to perform predefined actions on data packets matching the firewall rules, no matter how the firewall is deployed, the traffic that requires access control must be forwarded through the firewall. From the forwarding logic, the firewall always among the network traffic which means the network traffic will first enters the firewall and then leave the firewall. At present, most firewall deployment can be divided into routing, switch, transparent, policy routing and other modes. No matter which deployment mode the firewall is in, the relationship between data packets and the firewall must belongs to one or a combination of the following four models:

A. Normal model

In the normal model, both the forward and reverse data packets of the TCP packet will pass through the firewall and only pass once, and the inbound and outbound interfaces of the forward and reverse traffic are exactly opposite, as shown in Figure 1:

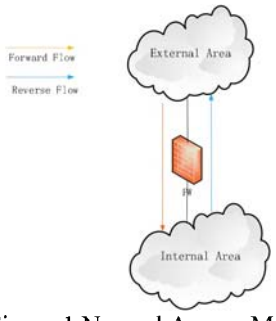

Figure 1 Normal Access Model

B. *One-sided model*

In the one-sided model, only forward or reverse network traffic passes through the firewall. Therefore, when the firewall wants to detect the flow status, the detection will fail which resulting in a network failure. The schematic diagram is as follows:

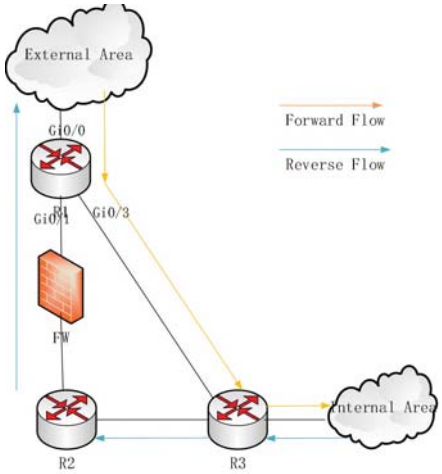

Figure 2 One-Sided Model

The common network configuration shown in Figure 2 is caused by poor consideration in early network planning. The line connection between R1 and R3 is used as the management connection for R1, and the gateway is set on R3. Some devices in the internal area are connected to the management network of R1 and these segments belong to the same network segment. Therefore, when the external needs to access the related equipment in the internal area, the route search is performed when the data message reaches R1. Because the local route is preferred, the data message is directly forwarded from R1 to R2, and then forwarded to the related equipment; When the reverse traffic returns to the external area, because the destination address don't belongs to the management network segment, it is forwarded along the path shown by the blue line in the figure, and passes through the firewall at this time.

C. *Round trip inconsistent model*

In the round-trip inconsistency model, both forward and reverse data packets will pass through the firewall, but the paths of the forward flow and the reverse flow are inconsistent. Because the flow session of the firewall is associated with the interface, the reverse flow cannot be connected to the forward flow. The association is established, which causes the firewall to function abnormally, as shown in Figure 3.

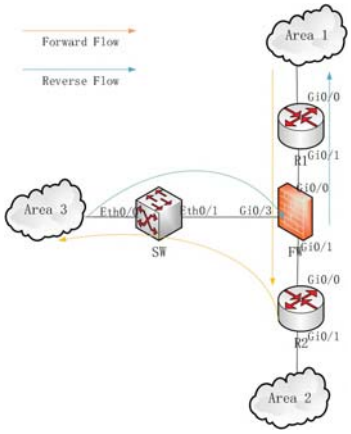

Figure 3 Round-trip Inconsistency Model

In the model shown in Figure 3, the gateway of the area device in Area3 is set on SW, the default route of SW points to R1.The routeing configured on R1 is that the destination address is Area3 and Area2 network segment, the next hop is R2. The routeing configured on R2 is as follows: the destination address is the Area3 device network segment, and the next hop is SW. In the actual environment, the configuration shown in Figure 3 may be very different, but finally they all lead to inconsistent round-trip paths.

D. *Repeat access model*

In the repeated access model, the forward and reverse data packets will also pass through the firewall many times, and the destination address of the data packets is the firewall, so this type of repeated access model is mainly aimed at the network management scenario of the firewall, as shown in Figure 4.

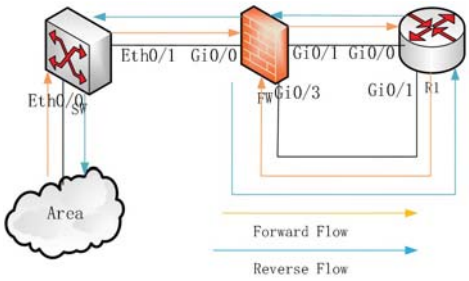

Figure 4 Repeated access model

In Figure 4, the deployment mode of firewall is switched mode or transparent mode, and the firewall's management gateway is set on R1. Therefore, when accessing the firewall from the area, the following situation will occur: first, the traffic is forwarded through the firewall to R1 on the level-2 layer, and then from R1 is forwarded to the firewall on the level-3 layer, that is, the traffic passes through the firewall twice and the destination address is the firewall.

In order to speed up the matching efficiency, firewalls are basically based on the rule: first packet matching of traffic, that is, the first packet will perform matching rule

search and matching process and cache the matching results for the flow. If the subsequent packets are the same flow, the cached results will be used directly. When the firewall performs subsequent matching, it will check the forward flow (for the case that the packet passes through the firewall multiple times, the path of the forward flow is based on the packet passing the firewall for the first time) and the reverse flow check. If: 1) Only forward or reverse traffic; 2) Asymmetric forward and reverse traffic paths. Both of these will cause traffic matching failures and traffic interruptions. Therefore, the one-sided model and the round-trip path inconsistent model will cause network traffic interruption, and this interruption often leads to unavailability of business. The repeated access model is the result of that the management port of the firewall cannot be isolated from the network space of the service port for some reason, which causes the same message to reach the firewall again, and thus cannot access the firewall management port. However, this situation often does not involve business. The harm is relatively minor.

2.2 DIAGNOSIS ALGORITHM FOR ANOMALY MODEL

This article abstracts various devices (firewalls, switches, routers, etc.) in the network topology into nodes in the graph, so the network topology can be abstracted as a undirected graph when judging the abnormal model.

Analysis in the one-sided model shows that the routing table of forward traffic on R1 shows that the next hop is R3, but the next hop of the reverse flow on R3 is not R1 but R2; in the round-trip inconsistency model, R1, R2, and SW are all connected to the firewall, and there is a routing for the same network segment on R2 and SW, and the next hop address of the route on R2 is SW, so further analysis can get: 1) When the routing in the topology is asymmetric and at least one of the devices at both ends of the routing is a firewall, the topology is an abnormal model; 2) When there are more than three devices connected to the firewall, and the destination network segment of three of the devices overlaps, and the next hop of the related routing of two of the devices is the firewall itself or one of the other two devices, it can be determined that the network topology is abnormal. Analyzing the repeated model, it can be concluded that when the mode of firewall is switching or transparent mode, and the traffic passes through the firewall and reaches the firewall again, there will be abnormal situations. However, this type of situation often does not affect the business and is easy to judge, so this article does not consider the detection algorithm for this model. In summary, an abnormal model detection algorithm can be derived, as shown in Figure 5.

```
struct list_head{
      struct list_head *next;
      struct list_head *prev;
}
struct routing_table{
     strtuct list_head head;
     __be32 dest_network;
     __be32 network_mask;
     __be32 next_hop;
}
struct node_network{
      char node_name[20];
      bool is_firewall;
      int number_neighbor;
      strtuct list_head route_head;
      __be32 ip_list[0];//ip list of interface
}
Graph network_graph;
bool Anomaly_model_detection( )
{
      struct node_network *next_hop=null;
      struct node_network *node_network=null;
      struct node_network *neighbor1,
*neighbor2,*neighbor3;
      bool is_normal=true;
      int firewall_included=0, status=0;
      struct routing_table *overlap_rout=null;
      list_for_each_graph_node(network_graph,node_net
work) {
      list_for_each_routing(node_network->routing_table
) {
            next_hop=find_next_hop(node_network->routin
g_table);
            If((!next_hop->is_firewall
&& !node_network->is_firewall) || !next_hop=null){
                  continue;//If both ends are not firewalls or
there is no next_hop device, the next routing is
processed}
            firewall_included++;
            status=0;
            list_for_each_routing(next_hop){
 If(dest_match(next_hop->routting,node_network)){
            If(routing_match(next_hop->routting,node_net
work)) {
                  status=1;
                  break;
               }
               }
            }
            }
            if(node_work->number_neighbor>=3 &&
node_work->is_firewall){//Detect inconsistent round-trip
models
            list_for_each_node_three_neighbor(node_work,
neighbor1,neighbor2,,neighbor3){
            overlap_rout=calcul_overlap_routing_and_abno
rmal(neighbor1,neighbor2,,neighbor3,node_work);
            If(!overlap_rout)
                return false;
            }
            }
      }
      if(status==0){
            is_normal=false;
            return is_normal;
      }
      if(firewall_included==0)
          is_normal=true;
      return is_normal;
}
```

Figure 5 Anomaly model detection algorithm

The algorithm shown in Figure 5 uses breadth-first search to traverse the graph, and processes each routing of each node in the graph: get the node corresponding to the next hop of the routing, and find out whether there is a routing for reverse traffic on the node. If the route does not exist, it indicates that there is an exception in the graph and returns false. If no firewall device is found after traversing all nodes or the routing involved in the firewall is always symmetrical with the peer end, the algorithm will return true.list_for_each_graph_node() is used to traverse the graph and returns each node of the graph; list_for_each_routing() is used to traverse a node and returns each routing of the node; find_next_hop() is used to find the node corresponding to the next hop of each routing; dest_match () is used to detect whether the next hop of the routing is the previous node; routing_match() is used to detect whether the routing of the previous node is symmetrical with the routing corresponding to the next hop node; calcul_overlap_routing_and_abnormal() is used to calculate the overlapping part of the routing of the nodes, and check whether there is an abnormality according to the above rules, and if there is an abnormality, it returns the destination network segment part of the overlapping route.Through simple analysis, it can be concluded that the time complexity of the algorithm is O(n+e), where e is the number of edges in the graph and n is the number of vertices in the graph.In this article, the graph is described in the form of adjacency linked list, and there is no recursive routing in the routing involved in this article, and complex network scenarios such as MPLS and VPN are not considered.

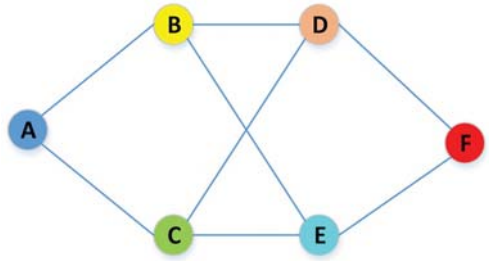

Figure 6 Typical network topology

Figure 6 is a schematic diagram of a typical network topology. If there is a route on B (firewall): the destination network 192.168.4.0/24, the next hop address is D; but the next hop is B first when the routing check is performed on the D node. Then check whether each routing can match the traffic with the source address network segment of 192.68.4.0. If they do not match, it can indicates that there is an abnormal situation.

III. EXPERIMENT AND ANALYSIS

This chapter will instantiate the network topology diagrams corresponding to the unilateral model, repeated model and abnormal access model proposed in the previous article, and give corresponding solutions. Of course, as mentioned above, there is more than one solution for each type of problem, and it is difficult to exhaustively list with personal thinking, so only one solution of the author is given for reference in this article.

3.1 One-Sided model

A.Key network configuration

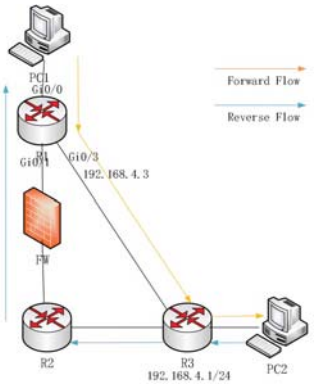

Figure 7 Example of One-Sided model

The detailed related network configuration is as follows:

TABLE I. CONFIGURATION OF ONE-SIDED MODE INSTANCE

Object name	Setting item	Setting content
PC1	Network configuration	IP address: 192.165.4.2/24 Gateway:192.168.5.1
PC2	Network configuration	IP address:192.168.4.2 Gateway:192.168.4.1
R1	Routing configuration	192.16.0.0 255.255.0.0 R2 address
	Interface configuration	IP address:192.168.4.3/24
R2	Routing configuration	192.168.4.0 255.255.255.0 R3 address 0.0.0.0 0.0.0.0 R1address
R3	Routing configuration	0.0.0.0 0.0.0.0 R2address

B.Solution

First of all, analysis shows that the reason why the forward traffic in Figure 7 exits from Gi0/3 is that the routing tables corresponding to Gi0/0 and Gi0/3 are in the same routing space, that is, the global routing table. Therefore, the new VRF is adopted in this article. Method to strip Gi0/3 from the global routing table, the specific steps are as follows:

1. Create a VRF on the router and name it mgt_vpn;

2. Add the Gi0/3 port of R1 to mgt_vpn and reset the IP address 192.168.4.3/24.

Perform related tests after completing the above steps, and the test results show that the communication between the terminals PC1 and PC2 is normal.

3.2 Round-trip inconsistent model

A.Key network configuration

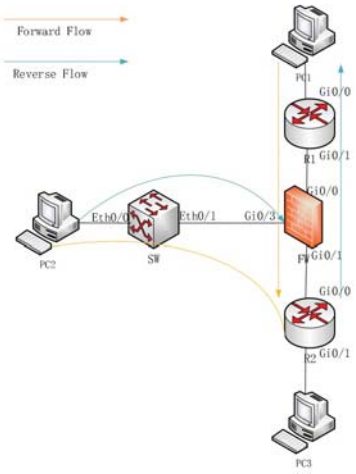

Figure 8 Example of round-trip inconsistent model

The detailed key network configuration is as flows:

TABLE II. CONFIGURATION OF ROUND TRIP INCONSISTENT MODEL INSTANCE

Object name	Setting item	Setting content
PC1	Network configuration	IP address:192.165.4.2/24 Gateway:192.168.5.1
PC2	Network configuration	IP address:192.168.4.2/24 Gateway:192.168.4.1
SW	Network configuration	IP address:192.168.4.1/24
	Routing configuration	0.0.0.0 0.0.0.0 R1 address
R1	Routing configuration	192.168.0.0 255.255.0.0 R2 address 192.168.4.0 255.255.255.0 SW address
R2	Routing configuration	0.0.0.0 0.0.0.0 R1 address 192.168.4.0 255.255.255.0 SW address

In the example shown in Figure 8, the switch and firewall are both configured in the switching mode, and the gateway address of the network segment 192.168.4.0/24 is set on the SW.

B.Solution

In this example, it can be seen that the traffic from PC1 to PC2 is first forwarded from R1 to R2 through the firewall due to improper routing settings, and then from R2 to PC2 through the firewall. It can be easily found that the method of using VRF for isolation here is no longer suitable. The solution here is: 1) Add a new routing on router R1: the destination address is 192.180.4.0/24, and the next hop address is SW; 2) delete the routing with destination address 192.168.4.0/24 on router R2. After add these changes, the results of testing show that the network communication between PC1 and PC2 is normal.

3.3 Repeated access model

A.Key network configuration

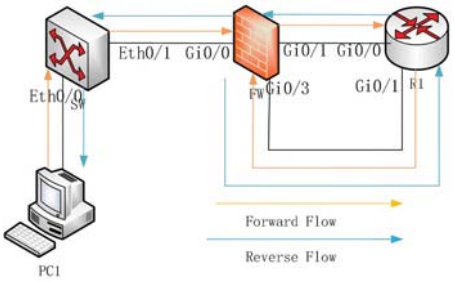

Figure 9 Example of repeated access model

The detailed key network configuration is as follows:

TABLE III. CONFIGURATION OF REPEATED ACCESS MODEL INSTANCE

Object name	Setting item	Setting content
PC1	Network Configuration	IP address :192.168.3.2/24 Gateway:192.168.3.1
Switch	Network Configuration	Configure the SVI port, and configure the corresponding IP address: 192.168.3.1
	Routing configuration	192.168.4.0 255.255.255.0 Address of R1
Firewal l	Interface configuration	Configure the IP address of Gi0/3:192.168.4.2 Gateway:192.168.4.1
Router	Gateway settings	IP address of Gi0/1:192.168.4.1
	Routing settings	0.0.0.0 0.0.0.0 Interconnection address of switches

The firewall in Figure 9 is configured with IP addresses for the management port Gi0/3, and the other two interfaces are configured as switching interfaces. Therefore, when PC1 accesses the firewall, data packets first pass through the firewall in the form of switch to the gateway address, that is, the router R1, then transferred out to the firewall through Gi0/3 on R1, which caused a network failure.

B.Solution

This kind of problem can be solved by the one of following methods: 1. Reconfigure the network and IP network segment, add a new SVI port on the firewall, use the address of the SVI port as the management address and the gateway is set on the SW, and close the Gi0/3 port of the firewall; 2. Create a new VRF and add The Gi0/3 port of the firewall is assigned to the newly-built VRF.

Method 1 has a large change, so this article chooses method 2 for experimentation. The relevant configuration steps can refer to the one-sided model of the previous article. The results of testing show that the network through the PC to access the firewall is normal after the modification

3.4 Verification of algorithm validity

In order to verify the effectiveness of the algorithm, this paper first abstracts the testing network topology of the one-sided model and the round-trip inconsistency model into a graph represented by an adjacency linked list similar to Figure 6, and represents each network device according to the data structure of the algorithm shown in Figure 5. Then, the routing table of each network device was filled with the structure representing the route according to the destination address, subnet mask, and next hop, and a simple verification was performed using the algorithm shown in Figure 5, and the algorithms all returned false. At the same time, replace the firewalls in the above test cases with router devices, that is, the algorithm returns true when the is_firewall item is false. The above experiments show that the detection algorithm can detect abnormal situations involving the firewall deployment mode in the network topology.

IV. SUMMARY

This article first briefly introduces the deployment mode of the firewall. It is clear that no matter which deployment mode traffic flows through the firewall, the characteristics can be described. On this basis, the normal model, unilateral model, repeated model and abnormal access model are proposed. Except for the normal model, the other models are abnormal models, which will cause abnormal network access. This article analyzes these abnormal phenomena and determines the root causes of the abnormalities: 1) The firewall detection mechanism needs to detect whether the forward and reverse traffic paths are symmetrical. 2) When the network space is consistent, the packets that repeatedly reach the firewall and the destination address is the firewall will be discarded. The basic idea of the solution is to ensure that the relevant forward and reverse traffic paths are completely symmetric or isolate the network space through technical means. At the same time, this article gives corresponding experiments and results in combination with examples, which has played a very enlightening role in solving related problems. On the basis of analyzing the causes of abnormalities, this article proposes an algorithm for detecting abnormal models. Experiments show that the algorithm is very effective in detecting abnormal models in a simple network environment. It is very useful for predicting the impact of firewall deployment modes and related network configurations on traffic.

REFERENCES

[1] Pankaj Gupta, Nick Mckeown. Algorithms for Packet Classification. IEEE Netwrok,2001.15(2):24-32;

[2] Dmifty Rovniagin, Avishai Wool. The Geometric Efficent Matching Algorithm for Firewalls. Proc. of the 23rd ;International Conference on Electrical and Electronics Engineers.2004,153-156;

[3] S.Singh, F.Baboesec, G.Varghese. Packet Classification Using Multidimensional Cuttings. SIGCOMM'03 Karlsruhe, Gemany:ACM,2003,213-224;

[4] V.Srinivasan, S.Suri, G.Vargheset. Packet Classification Using TupleSpace Search. ACM SIGCOMM Computer Computer; Communication Review. Oct 1999,29(4):135-146;

[5] Nitesh B. Guinde, Roberto Rojas-Cessa and Sotirios G. Ziavras. Packet Classification using Rule Caching. IISA 2013 , July 10,2013.

[6] N. Guinde, S.G. Ziavras and R. Rojas-Cessa, "Efficient Packet Classification on FPGAs also Targeting at Manageable Memory Consumption," in Proceedings of the 4th International Conference on Signal Processing and Communication Systems, Gold Coast, Australia, 2010, December 13-15, 2010.

[7] E.Al-Shaer, H.Hamed. Design and Implementation of Firewall Policy Advisor Tools. Tech.rep.,School of Computer Science Telecommunications and Information Systems, DePaul University,2004;

[8] M.Charalambides, P.Flegkas, G.Pavlou, A.K. Bandara, E.C. Lupu, A.Russo, N.Dulay, M.Sloman, J.Rubio-Loyola. Policy Confict Analysis for Quality of Service Management. Proc.of the Sixth IEEE International Workshop on Policies for Distributed Systems and Network.2005,99-108.