International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 02 Issue: 07 | Oct-2015 www.irjet.net p-ISSN: 2395-0072

Hashing Based Packet Matching Algorithm for Firewall

Ms. S. A. Dongre[1], Mr. S. G. Shikalpure[2]

[1]*Research Scholar, PG Student , CSE Dept, Government College of Engineering, Aurangabad, Maharashtra, India*

[2]*Assistant Professor , CSE Dept, Government College of Engineering, Aurangabad, Maharashtra. India*

Abstract - ***Firewall acts like sentry. It guards a corporate network by standing between the network and the outside world. So special attention must be paid to their packet matching algorithm which we are studying in this paper. This paper proposes an algorithm that is designed for divergence resolution and gives good network performance by reducing the packet matching time of the firewall. The proposed algorithm uses the method of hashing for matching the incoming packets with the main rule base. The performance of the algorithm has enhanced performance over other conventional algorithm in terms of packet matching time.***

Key Words: *Network security, Firewall, Packet filtering, hashing*

1. INTRODUCTION

Computer security is a hard problem. The dramatic rise and progress of the internet has opened possibilities that no one would have thought of. We can connect any computer in the world to any other computer, no matter how far the two are located from each other. This is undoubtedly a great advantage for individuals and corporate as well. However, this can be a nightmare for network support staff, which is left with a very difficult job of trying to protect the corporate networks from a variety of attacks. At the broad level, there two kinds of attacks:

- Most corporations have large amounts of valuable and confidential data in their networks leaking of this critical information to competitors can be a great setback.

- Apart from the danger of the insider information leaking out, there is a great danger of the outside elements (such as viruses and worms) entering a corporate network to create havoc.

As a result of these dangers, we must have mechanisms which can ensure that the inside information remains inside and also prevent the outsider attackers from entering inside a corporate network. Consequently, better schemes are desired to achieve protection from outside attacks. This is where a firewall comes into picture. See figure 1 which gives firewall with simple rules. Used properly, a firewall provides a significant increase in computer security.

The characteristics of a good firewall implementation can be described as follow[10],

- All traffic from inside to outside and vice versa, must pass through the firewall. To achieve this, all the access to the local network must first be physically blocked and access only via the firewall should be permitted.
- Only the traffic authorized as per the local security policy should be allowed to pass through.
- The firewall itself must be strong enough, so as to render attacks on it useless.

-Deny all inbound traffic with network address matching internal registered IP addresses
- Deny all inbound traffic to server from external addresses
- Deny all inbound ICMP echo request traffic

-Allow web traffic from any external address to the web server
-Allow traffic to FTP server
-Allow traffic to SMTP server
-Allow traffic to internal IMAP server

Firewall

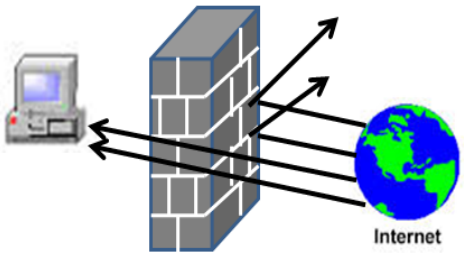

Fig. 1: Simple Firewall with Rules

2. MOTIVATION

The firewall is one of the central technologies allowing high level access control to organization networks. The scale of firewall rule set becomes larger, so average processing time required for each packet is also increasing. Hence performance and efficiency of the firewall get directly reduced. Thus, in this case the size of rule set becomes larger and larger. There are two ways to maintain the performance of the firewall:

1) Reducing the processing time required for each data packet matching with the rule base.
2) Reducing the size of rule set that each packet to match.

Packet matching in firewalls involves matching on many fields from the TCP and IP packet header. At least five fields are involved in the decision which rule applies to a given packet. That is

- protocol type(TCP/UDP)
- Source IP address
- Destination IP address
- Source port
- Destination port

With available bandwidth increasing rapidly, very efficient matching algorithms need to be deployed in modern firewalls to ensure that the firewall does not become a bottleneck.

Firewall packet matching is reminiscent of the well studied router packet matching problem [1]. However, there are several crucial differences which make the problems quite different. First, unlike firewalls, routers use "longest prefix match" semantics. Next, the firewall matching problem is four or five dimensional, where as router matching is usually one or two dimensional: A router typically matches only on IP addresses and does not look deeper, into the TCP or UDP packet headers. .Finally, major firewall vendor's support rules that utilize IP address ranges, in addition to subnets. Therefore, firewalls require their own special algorithms.

The firewall rules example are shown in table 1.and the format of the rules in the table upon the format used in Access Control Lists(ACL) on Cisco routers

Table -1: Sample Firewall Rules

Type	Source IP	Source port	Destination IP	Destination port	Action
TCP	1.2.3.1/5	1024	5.6.7.8	[1,65534]	accept
UDP	7.8.9.10	1025	11.12.13.*	90	refuse
TCP	11.12.13.*	*	20.21.*.*	*	refuse

3. LITERATURE REVIEW

3.1 Firewall Types

1. Packet Filter Firewall: These are based on first generation firewall technology. They analyze network traffic at the transport layer. They examine each IP network packet to see if it matches one of the rules defined for allowing or denying data flows. The decision is based on the information they get from the packet's transport layer headers and the direction the packet is going into. They are therefore configured to check:

Transport layer type (TCP, ICMP and UDP)
Source port
Destination IP address
Source IP address
Network interface the packet arrives on
Destination port

Packet filters do the above by applying a rule set residing in the TCP/IP kernel that defines what action goes with which rule[16].

2. Circuit Level Firewalls*:* These are based on second generation firewall technology. They work based on the fact that a packet is either a data packet or a connection request belonging to a connection or circuit between two peer transport layers. These firewalls work by:

- checking that each connection setup follows a handshake system for the transport layer protocol being used.
- Storing a session identifier for the connection - Connection state: handshake, established, or closing
- Only forwarding packets after the handshake is complete
- Maintaining a table of valid connections and removing it once the connection is terminated
- Closing the virtual circuit after transmission

3. Application Layer Firewall*:* Also called third generation firewall. These firewalls evaluate packets for valid data at the application layer before allowing a connection.

- Examines data in network packets at the application layer
- Maintains connection state and sequencing information
- Can validate passwords and service requests Most of them include proxy services for specific services such as HTTP or FTP which provide more checks and generate audit records about the traffic they transfer.

4. Dynamic Firewall: A fourth generation firewall type allowing modification of the rule base. A virtual connection is established and the packet is allowed to travel the firewall server. These provide support for UDP packets by associating them with a virtual connection. The connection information is kept for a short period and the connection is terminated if no response packet is received within that short time. They are good for not allowing unwanted UDP packets into a network because the

response packet must contain a destination address that matches the original source address.

5. Hybrid Firewall: Because of the need to do more than packet inspection, firewalls are being implemented as hybrid systems. These are mostly implemented by adding packet filtering to an application gateway. Cisco PIX firewalls are an example of such hybrid firewalls.

3.2 Related Work

Two additional categories of firewall exist depending on whether the firewall keeps track of the state of network connections or treats each packet in isolation.

1.**Stateful Firewall**: It deals with the state of connections, state here is defined as the condition of connection, which varies greatly depending on application or protocol used. In Stateful firewall when the first packet in a network is allowed to cross the firewall then all subsequent packets belonging to that flow and especially the return traffic flow is also allowed to pass through the firewall. Stateful firewalls typically build a state table and use this table to allow only returning traffic from connections currently listed in the state table. After a connection is removed from the state table, no traffic from the external device of this connection is permitted.
This statefulness has two advantages:

- No need to write explicit rules for return traffic and such return-traffic rules are inherently insecure since they rely on source-port filtering. This makes Stateful firewalls more secure as compare to stateless firewall.
- State lookup algorithms are typically simpler and faster than rule-match algorithms.

2. **Stateless Firewall:** In stateless firewall packet filters at network layer or it uses transport layer information only so they only look at the header part of a packet. The packet filter does not examine the data section of a packet. Action decides which service is to permits or denies i.e. to allow the packets or to drop them. Because of the stateless nature it needs to monitor all the incoming and outgoing packets which is time consuming as each and every packets need to be matched with the firewall rule list to check if the packets should be allowed or need to get drop out of the system. Also search mechanisms by a slow algorithm like linear search of the rule-base that implements the first match semantics makes its more time consuming.
Stateless Firewalls are the most basic and they are the most common type of firewalls. Stateless firewalls basically watch the traffic and compare the packets with the rules from its rules database. If a malicious activity is found it drops the packet. They are not aware of the traffic flowing among them. For simple lightweight host – based protections usually stateless firewalls are preferred. There are many examples for stateless firewalls: IP tables from Linux

4. RULE MATCHING ALGORITHM

Most modern firewalls are stateful. This means that after the first packet in a network flow is allowed to cross the firewall, all subsequent packets belonging to that flow, and especially the return traffic, is also allowed through the firewall.
Firewall statefulness is commonly implemented by two separate search mechanisms:

1. A slow algorithm that implements the first match semantics and compares a packet to all the rules, and
2. A fast state lookup mechanism that checks whether a packet belongs to an existing open flow.

4.1 Existing System

The firewall packet matching problem finds the first rule that matches a given packet on one or more fields from its header. Every rule consists of set of ranges [li , ri] for i = 1, . . . , d, where each range corresponds to the i-th field in a packet header. The field values are in 0 <=li ,ri<= Ui ,where $Ui=2^{32}$ - 1 for IP addresses,
Ui = 65535 for lport numbers, and Ui = 255 for ICMP message type or code.
The geometric efficient matching search data structure consists of three parts[1]. The first part is an array of pointers, one for each protocol number. The second part is a protocol database header, which contains information about the order of data structure levels. The order in which the fields of packet header are checked is encoded as a four tuple of field numbers.
The third part represents the levels of data structure themselves. Every level is a set of nodes where each node is an array. Each array cell specifies a simple range, and contains a pointer to the next level node. In the last level the simple range information contains the number of the winner rule instead of the pointer to the next level.
The search algorithm
The packet header contains the protocol number, source and destination address, and port number fields. First, we check the protocol field and go to the protocol array of the search data structure, to select the corresponding protocol database header. From this point, we apply a binary search with the corresponding field value on every level, in order to find the matching simple range and continue to the next level. The last level will give us with the desired result—the matching rule number.

4.2 Proposed work

Network traffic is increasing tremendously. Linear packet filtering takes more time to filter this enormous traffic,

firewall should be able to sustain a very high throughput, or risk becoming a bottleneck. Here we try to propose that efficient matching algorithm filters more number of packets in minimum time period hence time complexity required for matching packets with rule set is less.

In the proposed work hashing method is applied for lookup operation and at the same an index file is maintained. Index file consist of rule number from the main rule set and its respective hash values. For every captured packet, based on the header information its key value is computed and mapping is done against index file. If proper match is found it indicates that the particular packet with same header information has previously entered the network and so further lookup is performed on the log file and based on the decision field action is taken.

The whole matching process is shown in figure2.

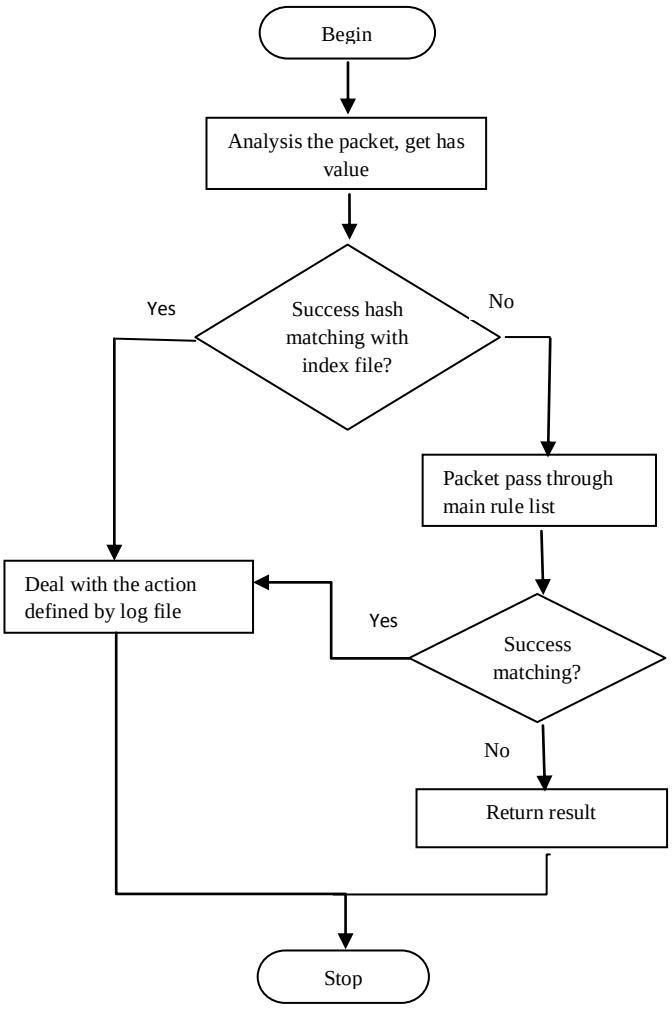

Fig.2. The flow chart of proposed matching algorithm

During implementation three files are maintained namely main file containing firewall rule, a log file which is subset of main rule set containing recently captured packets and an index file having hash values. For every captured packet, a key value is calculated based on its header information and mapping is done on the index file.

Initially the index file and log file is empty so for the first packet in the network flow lookup is performed on the log file and based on the decision field action is taken. On finding the exact match its hash value is computed and the corresponding entry is made in the index file and in the log file. All the succeeding packets belonging to the same flow performed matching by finding the record in log file instead of main rule set. Thus by cataloguing the information of the recently received packets we try to reduce the searching time to scan the main rule set. Here the log file is act as the subset of main firewall rule set. The number of rules in the log file is less as compared to the rules in the main file, so it is obvious that time required to scan the log file will be less as compared to time required for scanning the main rule set.

5. RESULTS

We have shown the results for number of packets matching time by using previous and proposed algorithmic approach with the help of graph as shown in figure 3.

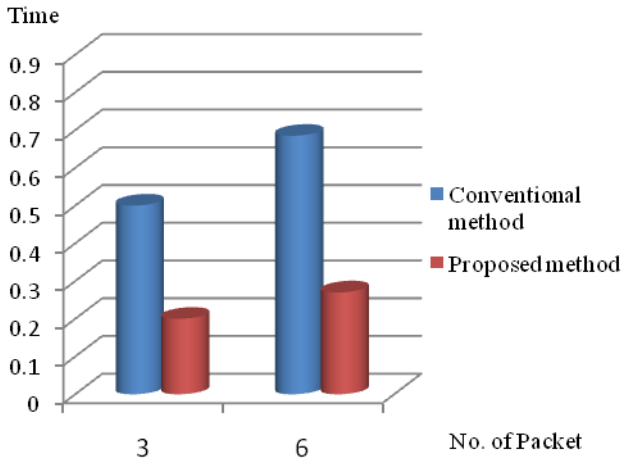

Fig. 3: Result Graph

6. CONCLUSION

A firewall packet filtering has become progressively important research in networking area over last few years. Many researches are making efforts to improve the rule policies and cost of matching needed. In this work by

managing the data in log file and an index file we are trying to reduce the time required for rule matching. Success definition of the work can be stated as average searching time needed for packet filtering and matching. We conclude that the proposed matching algorithm speed is faster than the linear search.

REFERENCES

[1] Dmitry Rovniagin and Avishai Wool, "The Geometric Efficient Matching Algorithm for Firewalls", Senior Member, IEEE, 2011.

[2] H.Bidgoli, Ed. JohnWiley& Sons, "packet filtering and stateful firewalls", in Handbook of Information security, 2006 Vol III

[3] "A quantitative study of firewall configuration errors", IEEE Computer, vol.37

[4] "Packet classification using tuple space search", , in Proc. ACM SIGCOMM,1999, PP135-146

[5] D. E. Taylor, "Survey and taxonomy of packet classification techniques", , ACMComput. Surv., vol. 37, no.3, pp, 238-275, 2005

[6] F.Baboescu and G. Varghese, "Scalable packet classification", in proc ACM SIGCOMM, 2001.

[7] M. G. Gouda and A.X. Liu, "A model of Stateful Firewalls and its Properties", , Proc, IEEE Conf. Dependable Systems and Networks (DSN 05), pp. 320-327, June 2005

[8] W.R. Cheswick, S.M. Bellovin, and A. Rubin, Repelling the Wily Hacker, "Firewalls and Internet Security:", second ed. Addison-Wesley, 2003.

[9] M. de Berg, M. van Kreveld, M. Overmars ,"ComputationalGeometry: Algorithms and Applications", 2nd ed. Springer-Verlag, 2000.

[10] Atul kahate, "Cryptography and Network security"

[11] Steven M. Bellovin and William R. Cheswick, "Network Firewall", IEEE Communication Magazine

[12] Avishai Wool,"Packet Filtering and Stateful Firewall",

[13] AlokTongaonkar, NiranjanInamdar and R. Sekar, "Inferring Higher level Policies from Firewall Rules"

[14] Shriya A. Jadhav and Dr. Pradeep K. Deshmukh, "Efficient Packet Filtering for stateful Firewall using the Geometric Efficient Matching Algorithm"

[15] Li Zhong and Li Xiao, "A performance- optimized firewall rules matching algorithm"

[16] Dr. P.K. Deshmukh, Dr. A. B. Bhagvan, Ms. P. Kinage, Ms. S. A. Jadhav ,"Investigation and Analysis of Efficient Firewall Packet Filtering and Matching Algorithms"