Computer Communications 32 (2009) 1893–1901

Contents lists available at ScienceDirect

Computer Communications

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

A TCAM-based solution for integrated traffic anomaly detection and policy filtering

Zhijun Wang [a,*], Hao Che [b], Jiannong Cao [a], Jingshan Wang [a]

[a] Department of Computing, The Hong Kong Polytechnic University, Hung Hum, Hong Kong
[b] Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX 76019, USA

ARTICLE INFO

Article history:
Received 12 November 2007
Received in revised form 20 June 2009
Accepted 17 July 2009
Available online 28 July 2009

Keywords:
Flow anomaly detection
Two-dimensional matching
TCAM coprocessor

ABSTRACT

The survivability of the future Internet is largely dependent on whether it will be able to successfully address both security and performance issues facing the Internet. On one hand, the Internet becomes more and more vulnerable due to fast spreading malicious attacks. On the other hand, it is under great stress to meet ever growing/changing application demands while having to sustain multi-gigabit forwarding performance. In this paper, we propose a Ternary Content Addressable Memory (TCAM) coprocessor based solution for high speed, integrated TCP flow anomaly detection and policy filtering. The attacking packets with spoofed source IP addresses are detected through two-dimensional (2D) matching. The key features of the solution are: (1) setting flag bits in TCAM action code to support various packet treatments; (2) managing TCP flow state in pair to do 2D matching. We evaluate the solution's ability to detect TCP-based flooding attacks based on real-world-trace simulations. The results show that the proposed solution can match up OC-192 line rate. The possible modifications of the solution for the detection of low rate TCP-targeted attacks are also discussed.

1. Introduction

The fast spreading malicious attacks make the Internet more and more vulnerable, while the ever growing application demands require more and more types of Internet services. The future Internet has to address both security and performance issues to survive.

Distributed Denial of Service (DDoS) attacks [11] are the major threats to the Internet. In DDoS attacks, attackers send a large amount of attacking packets using spoofed source IP addresses to a victim server which eventually runs out of its resources and degrades the performance of legitimate packets. One effective way to defend against DDoS attacks is IP traceback [3,6,9,16,17,20,21,23,24]. Using IP traceback, the attackers can be identified and punished through tracing their physical locations. However, effectively identifying attackers is difficult due to the stateless property of Internet routers/switchs. Hence it is important and challenging to do IP traceback.

The fast growing application demands need the network to provide various types of services. To support differential services, different packets may need to be treated differently based on, e.g., quality-of-service requirements or other policies. To this end, packet classification [2,15,18,25] based on a set of policy filtering rules must be performed in a router interface to identify the needed treatment of individual packets. Traditional policy filters [2,15,18,25] treat each packet individually, and does not attempt to associate the packet with other packets belonging to the same flow. Flow classification is a stateful packet classification, generally known as packet classification, which tracks the flow state by identifying every packet in every flow. Packet/flow classification has long been identified as the most critical data path function, creating potential bottlenecks for high speed packet forwarding.

To remove these potential bottlenecks, various algorithmic and hardware approaches have been developed, attempting to meet the targeted performance for efficiently PF. However, implementing different solutions to execute multiple tasks is very costly and sometimes even infeasible due to various resource constraints. Hence, it is both technologically and economically important to develop integrated solutions for PF and Content Filtering (CF), matching multi-gigabit line rate or even higher.

The traditional approach to enable security functions is generally separated from the approach that implements typical packet forwarding functions. For example, hash-based IP traceback [23] is generally implemented using dedicated chips for computing hash functions. Packet classification is typically performed as part of the packet forwarding functions in a router interface card, e.g., using a network processor and its associated Ternary Content Addressable Memory (TCAM) coprocessor [5,27]. A TCAM coprocessor contains self-addressable rules which map to different memory addresses in an associated memory (normally an SRAM) containing the corresponding actions. A rule matching in a TCAM is performed for all the rules in parallel. Each parallel matching is done at a time. Due to its high speed performance, the TCAM coprocessor is widely used as packet classifier in industry. How-

[*] Corresponding author. Tel.: +852 2766 7277.
E-mail address: cszjwang@comp.polyu.edu.hk (Z. Wang).

0140-3664/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2009.07.016

ever, the separated solutions add the complexity and integration costs to the next-generation Internet design. Hence, it is of both technological and economical importance to develop integrated solutions to enable security functions and high speed forwarding, matching multiple gigabit line rate.

In this paper, we propose a TCAM coprocessor based solution for high speed, integrated TCP traffic anomaly detection and policy filtering. The TCP-based DDoS attacks using spoofed source IP addresses are detected in the edge router through two-dimensional (2D) matching [8]. 2D matching means a normal TCP flow generated from one end host to another should have a corresponding flow from the other direction. The key features of the solution are: (1) setting flag bits in TCAM action code to support various packet treatments in the network processor and the local CPU; (2) managing TCP flow state in pair to do 2D matching. In the solution, when a TCP flow has not been matched after a period of time T_{alm}, the flow is considered to have high probability to be an attacking flow. Hence an alarm message composed of the flow identities is sent to the destination server, which in turn can use the information to do IP traceback. Based on the real Internet traffic analysis, the proposed solution requires about 5 Mbits TCAM memory to support OC-192 line rate for the integrated tasks. Such TCAM is available in today's market. We also discuss how to handle TCAM table overflow and analyze the solution's performance in case of table overflow. The simulations based on the real world traffic traces are conducted to evaluate the performance on the detection of TCP-based flooding attacks. The results show that the proposed solution can handle OC-192 line rate. The modifications on the proposed solution for the detection of low rate TCP-targeted attacks are also discussed.

The proposed solution focuses on the detection of TCP-based attacks.[1] The non-TCP-based attacks can be defended by other existing solutions such as Probabilistic Packet Marking [9,16,21,24] and Internet Control Protocol Message (ICMP)-based [3] IP traceback solutions. The real traffic measurement shows that the TCP traffic constitutes 80% of the total traffic (see Section 4), and the TCP-based attacks account for more than 80% of over all attacks [11]. Hence monitoring non-TCP packets for defense against DDoS attacks is much less challenging than monitoring TCP packets.

The rest of the paper is organized as follows: Section 2 describes the TCP traffic anomaly detection through 2D matching. The details of the integrated solution is presented in Section 3. The performance of the proposed solution is evaluated by simulations in Section 4. Section 5 discusses how to modify the solution to detect low rate TCP attacks. Section 6 briefly describes the related work. Finally, Section 7 concludes the paper and discusses some future work.

2. TCP traffic anomaly detection through 2D matching

In this section, we first give the needed definitions and then discuss how to detect anomalous TCP flows through two-dimensional (2D) matching.

A *flow* is a set of packets which have the same identity. The identity is extracted from the packet header. In this paper, the following five tuples: source IP address (SIP), destination IP address (DIP), source port number (SPN), destination port number (DPN), and protocol (PRO) are used as the *flow identity*. In other words, a flow is uniquely determined by the five tuples <SIP, DIP, SPN, DPN, PRO>.

TCP is a two-way communication protocol. A normal TCP flow generated from one end host (e.g., *A*) to another (e.g., *B*) should have a corresponding flow from the other direction (i.e., from *B* to *A*). Fig. 1 shows a general Internet architecture. Assume host *A* in Autonomous System 1 (AS1) sends a SYN packet to host *B* located in AS2 to initiate a TCP session. After receiving the SYN packet, host *B* sends a SYN + ACK packet back to host *A* to establish the session. In this case, the edge router 1 can detect both flows coming from AS1 (called *outbound flow*) and into AS1 (called *inbound flow*). For an outbound flow with flow identity <SIP, DIP, SPN, DPN, PRO>, the corresponding inbound flow identity is <DIP, SIP, DPN, SPN, PRO>. The feature of an outbound flow having a corresponding inbound flow is called two-dimensional (2D) matching [8]. An outbound (inbound) flow is called an *unmatched flow* if no corresponding inbound (outbound) flow arrives within a period of time T_{alm}. An inbound (outbound) flow is called the *matching flow* of its outbound (inbound) flow.

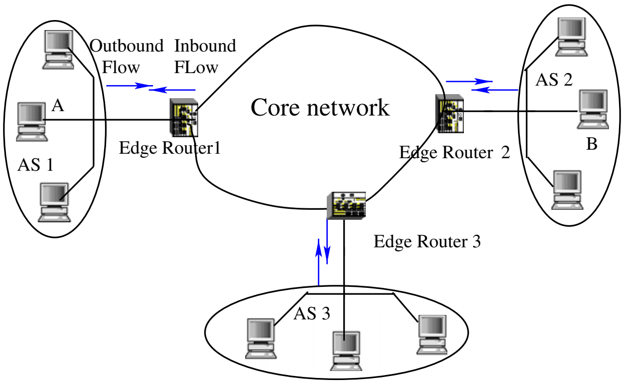

Fig. 1. Internet architecture.

2D matching can be effectively applied to detect TCP-based attacks using spoofed IP addresses. For attacking packets using spoofed source IP addresses, the responding packets are routed to the spoofed IP addresses which may be different from the original AS. Thus the edge router at the attackers' AS may only detect the outbound flow, and hence an unmatched flow is detected. Based on these observations, one can do 2D matching at the edge routers for TCP traffic anomaly detection. When an unmatched flow is detected, the router sends an alarm message (e.g., ICMP message) including the flow identity to the destination for possible IP traceback.

The most popular TCP-based DDoS attacks are TCP SYN and RESET flooding attacks. In these attacks, the SYN or RESET flag bit is set. To detect these attacks, we only need to maintain all the flows start with SYN and RESET packets. However, there are other types of TCP-based attacks [11] which have ACK bit set or no flag bit set. Hence, any TCP packet can be an attack packet. In our solution, if a packet does not belong to any existing flow, the packet is considered to be a new flow and will be monitored in the flow table to allow 2D matching.

Except the attacking packets, unmatched flows may be caused by: (1) the destination server is down; (2) the destination server has changed its IP address, but a cache entry of the old server IP address is still in the domain name server (DNS). In these cases, the destination is unreachable and the flows sent to the destination server can be viewed abnormally.

In the following sections, we will present the details on how to integrate 2D matching and policy filtering using TCAM coprocessors.

3. Integrated TCP traffic anomaly detection and packet classification

This section first gives a brief review of policy filtering using a network processor and its TCAM coprocessor, and then presents the details of the proposed solution.

[1] In fact, any attack with two-dimensional matching feature can be measured through the proposed solution. For example, ICMP echo reply attacks. For simplicity, we focus on the TCP-based attacks here.