A Study on the Anomalous Traffic Handling

KeunSoo Lee*, Sehun Kim**
Department of Industrial Engineering in KAIST
kslee@tmlab.kaist.ac.kr*, shkim@kaist.ac.kr**

Abstract

For recent years, hacking is in the trends of making excessive traffic unnecessarily to obstruct the service by getting a system's performance down. And we can also see systems paralyzed in service ability due to the flash crowds of normal traffic to a popular website. This is an important problem certainly solved in the point of QoS[1] guarantee for the clients. It is known that the former is DDoS(Distributed Denial of Service) attack and the latter is FE(Flash Event). These two are samely anomalous traffic because these make excessive congestion on the network or system and downgrade the system's service ability.

In this paper, we propose a scheme for protecting the system against anomalous traffic and for guaranteeing the QoS. For this, a server records and maintains the information of clients accessed more than one time before when it is in normal condition. When it falls into the congestion, the records are used for filtering the spoofed IP. We send and receive the ICMP request/reply packet to know whether the IP is spoofed. And we also propose for applying the object spliting of CDN[2] to guarantee the QoS in the initial FE situation.

1. Introduction

There are no other industrial fields made rapid growth more than IT. Internet provides much convenience and richness, so now more than about 70% of nation's popularity make use of the internet ,and the people can't spend a time without it. It is expected that the number of internet users will be increased because of the benefit and convenience the internet provides. But the rapid increase of users makes problem that we haven't seen before.

+ The work was sponsored in part by the Korean Ministry of Information and Communication in the context of University IT Research Center Project.

1 Quality of Service

2 Contents Delivery Network, The contents are provided from cache server near from the clients instead of origin server.

The problem is the excessive traffic generation on network or in a system, and these traffic gets the resources of a system exhausted. We call these kinds of hacking as DDoS attack when the traffic is generated by the malicious attacker and it is called as FE when the user is normal clients who want to get services from the system[2].

DDoS attack uses not the vulnerabilities of the system but the vulnerabilities of network protocol. There is no fundamental solution to defence or to traceback the source due to this characteristic.

We define these kinds of traffic generated by DDoS attack and FE as anomalous traffic, and propose a scheme for protecting a system efficiently against the anomalous traffic. And we propose a schems to guarantee QoS for the normal users as well.

This paper is organized as follows. We begin in Section 2 with discussion of the characteristics of DDoS and FE traffic. A scheme is proposed for the efficient handling of anomalous traffic in Sectioin 3, and the evaluation and analysis for the handling scheme is described in Sectioin 4. Finally, we conclude this paper in Section 5.

2. Characteristics of FE & DDoS attack

We described that the FE and DDoS traffic have same influences on network or a system with the exception of the source's intent mentioned at the above section. It was considered that it practically was impossible to distinguish these two traffic until these days by the reason of ultimately same characteristics between these traffic.

But some apparent characteristics which can tell who is who are presented in [4]. In the paper, the study was made in three dimensions, which are traffic pattern, clients characteristics and file reference characteristics.

First, there are no differences in the point that both generate unimaginable amounts of traffic in the dimension of traffic pattern.

Second, the flooding traffic is caused by the increase in the number of normal client's access to a system in FE situation, but in DDoS attack, it is the reason of rapid increase of malicious clients and unnecessarily excessive service request rate from the

lots of malicious agents[3]. Most of the clients who request service during FE made access more than one time before when the system was in normal condition, and the client distribution can be expected to follow population distribution among ISPs and networks, but lots of the clients who flood into a system during DDoS attack haven't accessed before and client distribution across ISPs and networks in DDoS attack doesn't follow the population distribution[4]. And the request rates of FE traffic are responsive to the performance level of destination server, but those of attack traffic are not responsive but stable. These are because that normal client's service requests are made after receiving the Ack packet from the distination, but the attack packets are transmitted in accordance with the time interval set by the attacker before not the reply acknowledgement.

Third, the service requests during FE are for some popular contents in a server, but in DDoS the requests don't concentrate on some particular contents.

3. Anomalous Traffic Handling

The characteristics for discrimination between FE and DDoS attack traffic are mentioned in Section 2.

A detection and handling scheme for the anomalous traffic is proposed in this section.

3.1 Anomalous Traffic Detection

The characteristic of rapid increase of request rates in short time length has been applied for anomalous traffic detection mostly.

For the operation of this, a server has to learn the traffic pattern while it is normal and must monitor and analyze the incoming traffic in realtime.

■ Normal Traffic Pattern Learning

The normal client's address is recorded at a memory called *NCID(Normal Clients IP Database)*. The information in the *NCID* is used for malicious packet filtering when anomalous traffic is detected, and also updated to perform more accurate filtering periodically. We apply the *sliding window* to eliminate the IP which is out of date and to keep the latest client's information.

[2] is introducing that the window size is 2 weeks and the information stored *NCID* is composed of IP address field and timestamp field used as basic information for the *NCID* update, but the window size can be resized in consideration of filtering level that you want and the cost of memory which is needed to record the client's information.

The old IP entries are removed when the value recorded in the timestamp filed exceeds the window size. We can save the cost of memory and can maintain the latest IP with the use of *sliding window* technology.

■ Traffic Monitoring and Analyzing

3 systems that generate attack packets to the target system. It is controlled by the real attacker througth secure network connection.

As you know, the request rates from the clients certainly make the queue size of a server changed. That is to say, when the request rates mark high, the queue size will be increased. And also it is very responsive to the rapidity of the request rates.

```
while(1){
  evaluate QueueWaitingTime ;
  if(QueueWaitingTime >= ThresholdWatingTime){
       then  determine AnomalousTraffic ;
             start HandlingProcess ;
             break ;
  }
}
```

Table 1 Anomalous traffic detection algorithm

The incoming process of packets in queue follows the poisson distribution. We can say that the waiting time, T_w, in queue is equal to the value that the queue size, N, is divided by the average number of incoming packets in unit time, λ. In general, T_w increases exponentially when the queue size, N, exceeds its tolerance level which is commonly called as threshold. So a server monitors the incoming traffic based on the request rates, evaluates its queue state, and continues to analyze the waiting time in queue. When the waiting time exceeds threshold, then handling processes are started. The frequency of traffic monitoring and analyzing processes should be taken in consideration of its performance in hardware and software for the minimization of overheads added to the server.

3.2 Spoofed IP Detection

For handling anomalous traffic, we apply the client distribution characteristic described in Section 2, and make use of the *NCID* for discrimination of spoofed IP from the traffic.

A security device, such as IDS or Firewall, performs the processes for overheads reduction on the server. The detail processes for detection of spoofed IP are described as follows.

(1)A target server sends a message requests handling and the *NCID* entry to the security device simultaneously.
(2)The source addresses of all incoming packets and the addresses in *NCID* are compared at the front security device.
(3)The spoofed IP is sorted out when both the IP address does not exist in the *NCID* and according as following processes.

①ICMP packet is sent to the source address that we want to see whether it is spoofed. The "Information Request/Reply" ICMP type can be used as you can see at Figure 1[1]. The packet should be sent with the records of suspicipus

IP in the 'Optional data' filed.

②The address in the reply packet's IP head and the value recorded in the 'Optional data' field before are compared. If these values does not match, the IP recorded in the 'Optional data' field is considered as spoofed IP, and the IP value is recorded in *SID*(*Spoofed IP Database*).

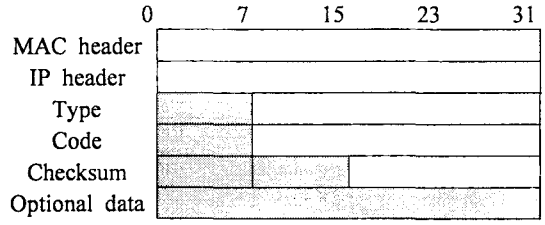

Figure 1 ICMP format

Needless to say, there are no any information in *SID* until the spoofed IP filtering processes are started. The *SID* is used for filtering the malicious packets, and the detail processes are explained at next subsection.

3.3 Spoofed IP Filtering

The *SID* doesn't store timestamp value like *NCID* but only a value related to the spoofed IP address because of its time information isn't necessary at the expense of processing and memory costs. With using *SID* we can provent the exhaustive repetitive spoofing test. The following describes the spoofed IP filtering processes.

(1)The source addresses of incoming packets are compared with the values in the *SID*. All packets are discarded when the source is exist in the *SID*. But when it isn't in the *SID*, then next process (2) is operated.

(2)The (3) process at the above subsection 3.2 is performed to see whether the IP is spoofed.

We can see all the handling processes for detecting and handling anomalous traffic at Figure 2.

By composing the *NCID* and *SID*, the flitering can be done more accurately, but we can't avoid the cost of memory. And the fast handling of incoming traffic is indispensable for realtime processing and overheads reduction. We recommend a hash technique for the fast IP lookup and to save the memory cost. Actually [2] insists that the memory costs can be saved to the minimum value which is no more sensive.

3.4 FE Traffic Handling

The flooding traffic concentrates on a origin server can be distributed, and also we can handle the serious congestion problem caused by FE traffic by applying CDN. But it is inevitable to experience fearful congestion when FE is in the initial situation.

We propose to apply the *object spliting* technology in CDN to get over congestion caused by the initial FE traffic.

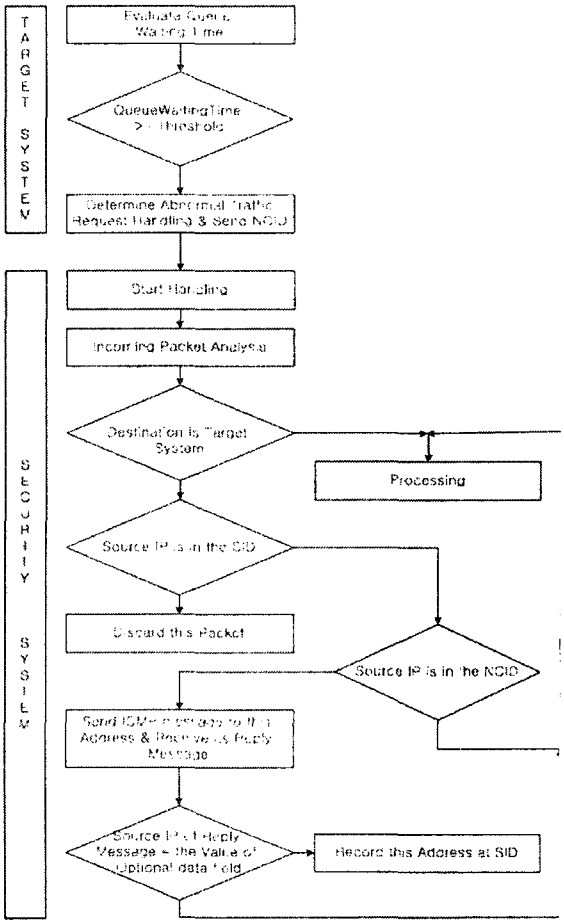

Figure 2 Handling Flowchart

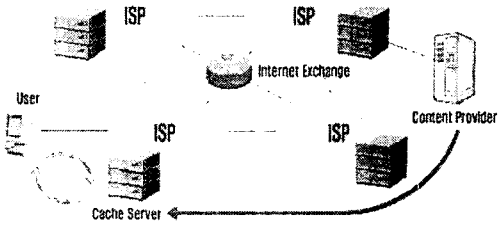

Figure 3 CDN Service

As you can see at Figure 4, the cache server only requests the contents which are requested from most of the clients to the origin server not missing most of the client's requests to the origin. The clients can get what he(or she) requested to the origin server from the cache server after caching the contents to it. If we apply and install the *object spliting* technology to all the networks and systems, then we can reduce the overheads imposed on the network or a system to about more than many times.

Nowadays CDN is in the way of expansion among the big web contents provider, and it is not difficult to perform the *object spliting* if the CDN is

applied more widely than now.

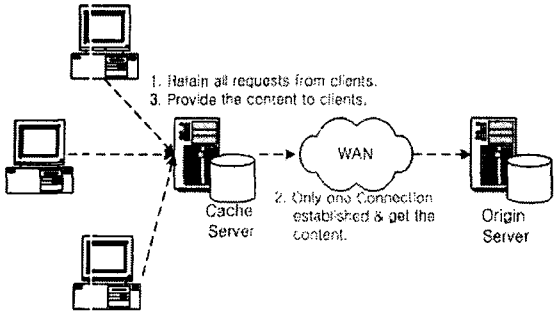

Figure 4 Object Spliting

4. Evaluation and Analysis

We described schemes for anomalous traffic handling at the above Section 3, and now analyze its charateristics, strengths and weaknesses at the following; paragraph.

First, the characteristic of rapid increase in request rate is applied to detect the anomalous traffic. And it is the common characteristic in two traffic, so we can expect to detect more accurately.

Second, the spoofed IP filtering is performed based on the client's access information before. To record the access information of clients, we need additional memory costs to compose, that is called, *NCID* and *SID*.

Third, we have to put up with the overheads caused by the handling processes, such as ICMP packet sending, comparing with each other to confirm the IP, monitoring realtime traffic, and analyzing *NCID* and *SID*.

Fourth, we can expect an additional effect for the attacker not to spoof the IP as a result of the smart filtering based on the ICMP reply information.

Finally, the *object spliting* technology in CDN makes us possible to avoid the serious congestion on the origin server during the initial FE situation because the service requests are distributed to each cache server which is near from the clients. And the origin server can provide high quality of service to its clients.

5. Conclusion and Future Study

Until now, we proposed schemes to handle the anomalous traffic caused by FE and DDoS attack. The malicious packets such as the spoofed IP should be sorted out from the normal traffic to quarantee high quality of service and to protect the target server. In the view of this point, a more intelligent spoofed IP filtering scheme is proposed in spite of overhead costs. And to provide the continuous services and prevent the preformance degradation of the origin server in the initial FE situation, the *object spliting* in CDN is proposed to apply to network. If we operate the proposed schemes, we can no more get in trouble with the anomalous traffic such as FE and DDoS traffic.

The Internet Users will continue to complain about the QoS problem as long as these kinds of network problems don't disappear on the network, and so it must be solved inevitably.

To do this, we have to study on more accurate malicious IP filtering and load balancing problems continuously for the future.

References

[1] 윤종호, "TCP/IP 및 윈도우 네트워킹 프로토콜", (주)교학사, 1999.

[2] CDNetworks, "http://www.cdnetworks.co.kr/cdn/internet_cdn.html"

[3] Seungrak Choi, Cheulung Yang, Jungsik Lee, "Core component technologies and trends of CDN", KISS, 2002.

[4] Jaeyoung Jung, Balachander Krishnamurthy, Michel Rabinovich, "Flash Crowds and Denial of Service Attacks:Characterization and Implications for CDNs and Web Sites", WWW10, WWW2002, May 7-11, 2002

[5] Tao Peng, Christopher Leckie, Kotagiri Ramamohanarao, "Protection from Distributed Denial of Service Attack Using History-based IP Filtering", IEEE2003 International Conference on Communications, 2003.

[6] William Stallings, "High-Speed Networks TCP/IP and ATM Design Principles", Prentice Hall, 1998.