NetFlow Based System for NAT Detection

Vojtech Krmicek
Faculty of Informatics
Masaryk University
Botanicka 68a, 602 00, Brno
Czech Republic
xkrmicek@fi.muni.cz

Jan Vykopal
Institute of Computer Science
Masaryk University
Botanicka 68a, 602 00, Brno
Czech Republic
vykopal@ics.muni.cz

Radek Krejci
Faculty of Informatics
Masaryk University
Botanicka 68a, 602 00, Brno
Czech Republic
kacak.r@mail.muni.cz

ABSTRACT

Revealing the misuse of network resources is one of the important fields in the network security, especially for the network administrators. One of them is the use of unauthorized NAT (Network Address Translation) devices (e.g. small office routers or wireless access points) inside the network which introduces serious security issues. There are several techniques proposed on how to detect NAT devices in the computer networks, but all these methods suffer from high false positive rate. Also there is no study how to perform NAT detection using NetFlow data, often used for monitoring and forensics analysis in large networks. The contribution of our work consists of the following: i) we have transformed existing NAT detection techniques to work with NetFlow data, ii) we propose three new NAT detection approaches, iii) we have designed a prototype of NAT detection system, which aggregates the results from various NAT detection techniques in order to minimize false positive and false negative rates.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*security and protection*

General Terms

Security

Keywords

Network address translation, NetFlow, detection, network security

1. INTRODUCTION

Network Address Translation is a procedure in which a special device, e.g. router performs modifications to the IP address and often also port number of a packet and maps the IP address from one space to another (usually from private IP addresses to public IP addresses and vice versa). This procedure can be easily misused because it is possible to hide more computers behind single IP address without the knowledge of the system administrator. Such behavior can be used e.g. for stealing the Internet connectivity or breaking ISP restrictions.

CoNEXT Student Workshop'09, December 1, 2009, Rome, Italy.
ACM 978-1-60558-751-6/09/12.

NetFlow traffic monitoring [2] is widely used approach for monitoring and forensics analysis in large networks. One of the major differences of NetFlow as compared to a traditional inspection method is that there is no packet payload information in the flow field. An ordinary flow record does not carry any high-layer information, it just contains traffic statistics. Therefore it is not possible to use NAT detection methods inspecting packet payload in order to detect NAT device.

To overcome these limitations we have extended the flow record with three new fields: TTL field, IP ID and TCP SYN packet size field. We extract the 8 bit TTL value, the 16 bit IP ID value and 16-bit TCP SYN value from the first packet belonging to new flow, flow is created and added to the flow cache. These extensions to the standard NetFlow format are then used for NAT detection purposes.

In the next section, we describe two existing NAT detection methods and we propose another three NAT detection techniques. An architecture of the whole NAT detection system with short description of its layers is introduced in Section 3. We conclude this abstract with the ideas of our future work.

2. NAT DETECTION METHODS

This section describes five proposed methods for NAT detection using NetFlow data. Two of them (**ΔTTL** and **IP ID** method) were previously published and we have modified them to use NetFlow data as an input. Furthermore, we propose other three methods for NAT detection based on NetFlow data.

- **ΔTTL** method is based on the approach described in [4] and detects NAT devices by identification of different TTL values introduced by various operating systems behind the NAT device.
- **IP ID** method is based on the approach described in [1]. It identifies IP ID sequences in the communication from particular host revealing possible NAT device.
- **ΔSubnet TTL** method performs analysis of TTL distribution over the inspected subnet and identifies NAT devices by the analysis of different TTL values then expected in the observed subnet.
- **ΔTCP SYN** method detects NAT devices by the identification of various length of TCP SYN packet introduced by various operating systems behind the NAT device (passive OS fingerprinting).

- **Port sequences** method analyses various port sequences in the communication from a particular host which can indicate more hosts hidden behind the NAT device.

All these methods are based on passive detection approach, meaning that we just observe network traffic. There is a possibility to improve detection methods and the whole detection accuracy by incorporating active detection methods, but we leave this active detection for future work experiments.

3. SYSTEM ARCHITECTURE

The architecture of our proposed prototype system has four layers where each layer processes the output of the layer underneath as shown in Figure 1.

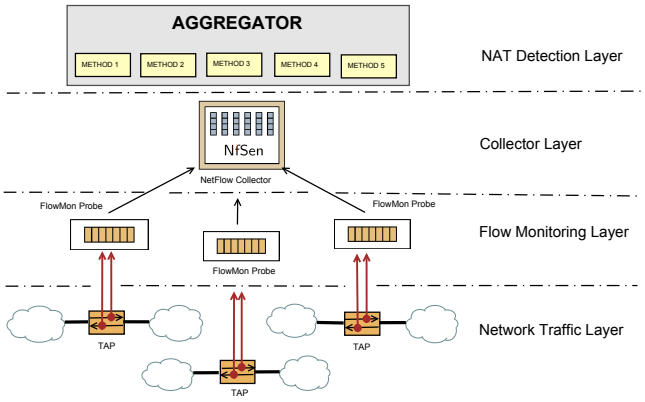

Figure 1: System overview.

Network traffic layer provides the data from the production network using the TAP (Test Access Port) devices. TAPs create real-time copy of all data on the observed links. Using TAPs avoids various traffic trace artifacts due to monitoring via port mirroring [5]. Non-intrusive observation allows deployment in networks with high availability requirements and "the most interesting" data.

Flow monitoring layer acquires the data from the network using the NetFlow probes (we use FlowMon probes [3]). The probes generate NetFlow statistics extended with the following items: TTL, IP ID and TCP SYN packet size. Flow statistics are sent via NetFlow version 9 protocol to the remote collector.

Collector layer principally consists of collector tools which are able to handle output from several NetFlow probes. Flow data are stored to a flat file database, including extended flow record information and are ready for further NAT detection.

NAT detection layer contains NetFlow based detection methods (see Section 2). The methods are implemented as a set of detection scripts. In order to reduce false positive rate we aggregate the output of individual methods using the NAT aggregation algorithm. This algorithm estimates for each host the probability of being the NAT device and presents this result to the network operator.

4. FUTURE WORK

In our future work we would like to focus on the following problems. Firstly, we will perform the detailed evaluation of the whole system by series of experiments. Namely, we will use the laboratory environment for simulating various NAT devices in the network and also we will evaluate the system using the real university campus traffic.

Other effort will be spent on the improvement of the methods to be able to work in heterogeneous environment, improve their false positive rates and especially focus on the more advanced method for aggregating their results.

Very promising approach is to incorporate active detection methods. When employing active detection, we would have possibility to discover network topology, perform active host fingerprinting etc. and these measurements would improve overall system results.

5. ACKNOWLEDGMENTS

This material is based upon work supported by Czech Ministry of Defense under Contract No. OVMASUN200801.

6. REFERENCES

[1] S. M. Bellovin. A technique for counting natted hosts. In *IMW '02: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 267–272, New York, NY, USA, 2002. ACM.

[2] Cisco Systems. Cisco IOS NetFlow. http://www.cisco.com/go/netflow, 2009.

[3] INVEA-TECH. Standard FlowMon Probe. http://www.invea-tech.com/, 2009.

[4] Peter Phaal. Detecting NAT Devices using sFlow, http://www.sflow.org/detectNAT/, 2003.

[5] J. Zhang and A. W. Moore. Traffic trace artifacts due to monitoring via port mirroring. In *Proceedings of the Fifth IEEE/IFIP E2EMON*, pages 1–8, 2007.