NAS Honeypot Technology Based on Attack Chain

Bing Liu, Hui Shu(✉), and Fei Kang

State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China
shuhui123@126.com

Abstract. With the wide application of network attached storage (NAS), the security problem is becoming more and more serious, together with the increasingly fierce attacks against NAS devices of different manufacturers. In order to better capture various types of attacks against NAS devices and detect security threats in time, this paper proposes a solution named NAS honeypot, mining the potential security threats through modeling and analyzing the NAS threats. Then a NAS honeypot based on device virtualization technology and depth monitoring and attack induction was designed on the basis of the construction NAS attack chain. Experimental results show that the NAS honeypot can effectively capture, record and analyze network attacks against multi-types NAS devices with a strong guiding effect in mastering popular attack method of NAS devices and alleviating the security threats of NAS.

Keywords: NAS · Threat modeling · Attack chain · Honeypot

1 Introduction

In recent years, with the rapid development of network technology and improvement of computer data processing capability, information and data on the network have been increasing explosive, bringing unprecedented demands to the data access, transmission and processing capacity of storage system. NAS has emerged with the increasing demand for storage capacity and reliability. Taking data as the center, it completely separates the storage device from the server and connects a group of computers by standard network topology [1]; clients are allowed to access data directly on the NAS without the server, providing a solution for heterogeneous platforms to use unified storage system [2].

Compared with traditional storage, NAS shows strong advantages, but it also causes new security problems. NAS not only has to assume the threat risk of traditional storage system, but also has the security risks from the network, such as data leakage caused by network transmission. Data resources are the most valuable in the Internet age, but security problems frequently emerge in NAS, a direct storage of data, in recent years. Nearly all manufacturers' devices have not been spared. CVE disclosed more than 300 related public vulnerabilities and almost covered all types. For example, for Gain Privileges: CVE-2018-18471, attackers can execute remote commands as root on the specific NAS of Seagate and NETGEAR [3]; for Bypass Something: CVE-2016-10108,

Q. Liu et al. (Eds.): CENet 2020, AISC 1274, pp. 915–926, 2021.
https://doi.org/10.1007/978-981-15-8462-6_106

Western Digital MyCloud NAS allows an unauthenticated attacker to run remote command as root [4]; for Code Execution: CVE-2020-9054, Multiple ZyXEL NAS devices allow a remote, unauthenticated attacker to execute arbitrary code [5]; for Gain Information: CVE-2018-13291, Synology DiskStation Manager allows remote authenticated users to obtain sensitive information [6]; for Directory Traversal: CVE-2018-13322, Buffalo TS5600D1206 allows attackers to list directory contents [7]; for Overflow: CVE-2018-14749, buffer overflow vulnerabilities in QTS can have unspecified impact on the NAS [8] and so on.

To address the endless security problems of NAS, we carry out research on NAS honeypot technology to capture attacks and enhance the security of NAS. Honeypot technology enables the defense side to clearly understand the security threats they face, and enhances the security protection ability of the actual system by means of technology and management, so it has been extensively and deeply studied in the field of network security. Moore [9] applied the honeypot technology to ransom worm detection, using two services to control the Windows security log, and establishing a graded solution strategy for attack; Saud et al. [10] used NIDS and KFSensor honeypot to detect APT attacks actively, and sent alarm information to the console when the honeypot service is called and run upon request. Anirudh et al. [11] proposed a solution of DoS attack honeypot for IoT devices, using IDS intrusion detection system to process client requests, and comparing information with log library to isolate and guide abnormal requests to honeypot and record abnormal source information. With the continuous development of honeypot technology, it is applied more and more in different scenarios and realized different functions. However, the research on NAS honeypot technology needs to be carried studied further.

In this paper, security threats are comprehensively analyzed and NAS attack chain is constructed through NAS threat modeling. NAS honeypot which is based on virtualization technology, interact with the attacker deeply through the virtual response to the attack, to induce them to launch in-depth attacks, capture and analyze the techniques and avenues of attacks. This paper solves the common NAS honeypot's problems of poor flexibility, poor expansibility and lack of unified deployment and control mechanism. The prototype system is implemented and the effectiveness of the proposed scheme is verified by actual deployment and experiment.

The rest of this paper is as follows: the second part introduces the NAS threat model and attack chain; the third part introduces the design and framework of NAS honeypot; the fourth part describes the experimental test; the fifth part summarizes the whole paper and looks forward to the next step.

2 NAS Threat Model and Attack Chain

The honeypot, which is designed based on the attack chain, can classify captured attacks according to specific attack chains, and intuitively and accurately identify the methods of attack is conducive to targeted security measures. We analyze and identify a variety of threats and separate attack steps of different attacks though threat modeling, and finally form a general attack link that can cover all attacks.

2.1 NAS Threat Model

Threat modeling is simply a structured method for identifying, quantifying and responding to threats. It helps to think about risks through abstract methods after the software design stage and before the software deployment stage. Generally, the process of threat modeling can be divided into 6 steps: identifying assets, creating an architecture overview, decomposing the application, identifying the threats, documenting the threats, and rating the threats [12]. We use the six-stage process to carry out threat modeling analysis for NAS.

First, we identify the assets that need to be protected by decomposing the composition of NAS hardware and software, including NAS devices, firmware, system application, mobile application and data storage, etc. By refining asset identification, we can discover as many types and numbers of threats as possible.

Then, we analyze the architecture, physical deployment, application functions and related technologies of the NAS system, so as to find potential vulnerabilities in the design or implementation of the application. We summarize the system architecture by creating the NAS system architecture diagram, and further identify the trust boundaries of the system, identify the data flow in the system, form the NAS system architecture and data flow diagram (Fig. 1), and then identify the entry points, and focus on the entry points and the data flow across the trust boundary, because the security of transferring data from outside is not trusted. We should focus on the threat analysis of this kind of data. According to Fig. 1, we can confirm entry points that affect system security, including firmware, system application, mobile application, user client, router, cloud and so on.

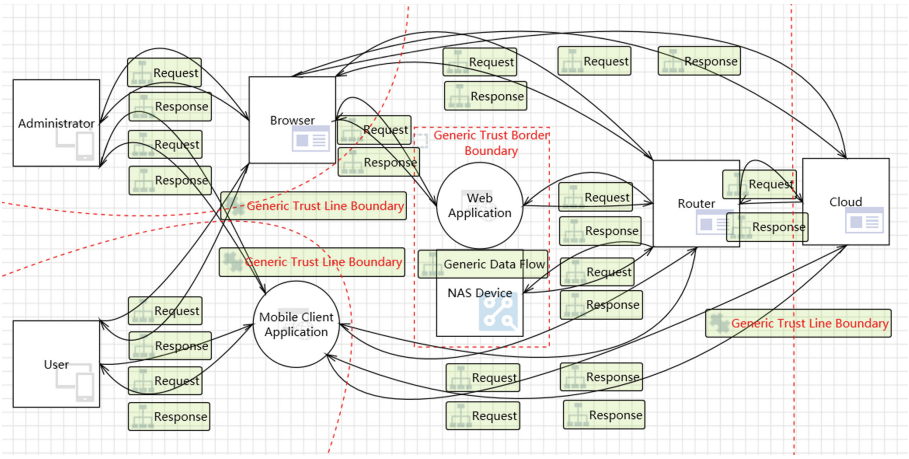

Fig. 1. The architecture and data flow of NAS system.

After that, based on the previous work, the STRIDE model is used to identify threats to NAS devices. STRIDE derived from an acronym for the following six threat categories: Spoofing, Tampering, Repudiation, Information disclosure, Denial of

service, and Elevation of privilege. We classify and identify potential security threats according to these six threat categories, such as obtaining the account password of device login, tampering with device configuration, tampering with device log record, obtaining device sensitive information, denial of service attack, remote control of NAS device, etc. Finally, we take a record of threats, take "Remote control of NAS device" as an example (shown in Table 1).

Table 1. The threat record of "Remote control of NAS device".

Threat description	Remote control of NAS device
Threat target	NAS user, NAS communication data, NAS device
Attack techniques	Improper use by NAS users, failure to modify the original password in time, or disclosure of the password in other ways; Attackers can analyze and find the administrator password by sniffing and intercepting the NAS communication data; The control of the device can be obtained by means of brute force, the use of firmware vulnerabilities and bypassing authentication
Countermeasures	Do well in guiding the use of the user's equipment, do well in the password protection of equipment, regularly change the password of equipment; Do well in auditing communication data, do well in communication data encryption; Set up multiple password error locking; Control firmware permissions strictly, and audit the authentication authority

2.2 NAS Attack Chain

NAS attack chain refers to the complete attack path that is composed of a general description of each attack stage of NAS and the attack sequence. Detecting and discovering devices is the first step to achieve NAS attacks. Therefore, the first part of NAS attack chain is NAS Detection. Then, we complete the attack chain through comprehensive analysis of threats identified by threat modeling. The representative high-risk threats are selected and the attack path is symbolized with STRIDE. To facilitate the description, NAS Detection is denoted as DT. Some symbolic descriptions of high-risk threats are shown in Table 2.

Through the symbolic description of threat, we can find that when NAS is discovered, some attackers will choose Spoofing attack (S) or Elevation of Privilege attack (E), and some attackers will choose to bypass the authentication attacks. We define the attack stage of S or E as Penetration Attack, as the second part of the attack chain. If achieving penetration attack or bypassing the authentication successfully, the attacker can carry out further operations. This stage is defined as the third part of the attack chain: Attack and Invade. When an attacker invade the device successfully, he can carry out various malicious attacks on the device. This stage is defined as the fourth part of the attack chain: Malicious Control. NAS attack chain is formed as follows:

Table 2. Symbolic description of threat.

Threat	Description
Tamper with device configuration	DT → T/R DT → S→T/R DT → E→T/R
Denial of service attack	DT → D/R DT → S→D/R DT → E→D/R
Remote control of NAS	DT → S→T/R/I/D/E DT → S→E → T/R/I/D DT → E→S/T/R/I/D

① NAS Detection → ② Penetration Attack → ③ Attack and Invade → ④ Malicious Control, the second stage can be skipped.

3 Design and Framework of NAS Honeypot

Designing honeypot by NAS attack chain can accurately grasp the stage of interaction with attackers, timely respond to attackers' attack actions, improve the probability of attackers attacking the honeypot, and lure attackers to conduct in-depth attacks.

3.1 Design of NAS Honeypot

The first part of the attack chain is NAS Detection. Attackers usually try to find NAS devices through search engines, port scanning, ICMP detection packets and so on, and determine the manufacturers and models of NAS devices according to the fingerprint information of equipment. NAS devices of some manufacturers will leak the firmware versions of devices and other more detailed information. In order to induce attackers to carry out further targeted attacks, the information is chosen back to the attacker in time by virtual response mechanism.

The second part of the attack chain is Penetration Attack. After attacker discovers NAS, it needs a highly simulated NAS honeypot to interact with him. In order to respond to attacks by different attackers against different honeypots, it needs to simulate NAS devices of different manufacturers and models based on virtualization technology. Therefore, it is necessary to establish virtual NAS sets, including NAS honeypots of different manufacturers.

The third part of the attack chain is Attack and Invade. In view of the attack of the attacker at this stage, we mainly consider providing support for subsequent attack classification, and set up specific attack detection units in each NAS honeypot, such as firmware update, configuration tamper, buffer overflow, etc., and the corresponding detection unit can be added according to the characteristics of equipment and detection needs.

The fourth part of the attack chain is Malicious Control. When capturing the attacker's malicious behaviors successfully, how to determine which are known attacks, which are unknown attacks, which are attacked and expanded, and how to record these attacks for better research and analysis. To solve these problems, we need to identify and classify attacks.

3.2 The Framework of NAS Honeypot

According to the design idea in 3.1, we designed and implemented the NAS honeypot. Its architecture is shown in Fig. 2. It including 5 parts: Control Center, Data Center, Virtual Response Center, Virtual NAS Set, and Attack Recognition Center. Data Center mainly collects attack data, classifies and collects the attack data collected by each part, so as to follow-up research and analysis.

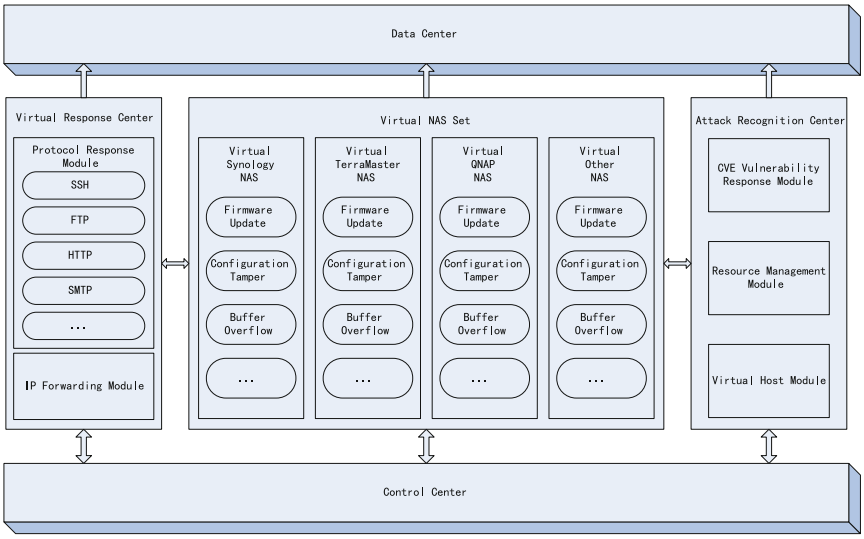

Fig. 2. The framework of NAS honeypot.

Control Center. The control center mainly controls each part of honeypot to respond to the attacker's attack, and connects to an external fingerprint database. In fact, the external fingerprint database is a collection of response packets for each vendor and each protocol. By sending different protocol query packets to different vendor devices, the response packets are extracted, classified and sorted out to form a fingerprint database for response packets for NAS attack detection. When receiving the protocol and port information sent by the Virtual Response Center, Control Center randomly selects the corresponding response packets from the external fingerprint database, returns it to the Virtual Response Center, controls the Virtual NAS Set to activate the corresponding virtual NAS, transmits the network data to the virtual NAS, and control the Virtual Host Module of the Attack Recognition Center connecting to the virtual

NAS. Record the attacker's attack behavior in the virtual NAS, and interact with the Attack Recognition Center to form a complete attack path.

Virtual Response Center. The Virtual Response Center mainly implements the response to the attacker's probe packets, including the Protocol Response Module and the IP Forwarding Module. The Protocol Response Module sets up a variety of protocol response units (such as SSH, FTP, HTTP, SMTP and other protocol types, which can be expanded according to actual needs). When receives the attacker's probe packets, the Virtual Response Center identifies the type of packets by feature extraction. The corresponding protocol response unit transfers the protocol and port information to the Control Center, and sends the response packets generated by Control Center to the attacker. IP forwarding module records the IP information of the attacker, and forwards the IP to the Virtual NAS Set after the virtual NAS is started, so as to deceive the attacker for subsequent attack. The Virtual Response Center sends attacker's detection behavior data to the data center.

Virtual NAS Set. The virtual NAS set is made up of virtual NAS from different manufacturers. When one of the virtual NAS is successfully launched, it becomes a specific NAS honeypot. Based on FIRMADYNE [13], we realize the simulation of every virtual NAS in the Virtual NAS Set.

First of all, we need to obtain firmware for NAS devices. Most NAS vendors constantly upgrade firmware to improve the function and security of the system, and provide firmware downloads on their official website, so we can get the firmware of mainstream NAS vendors through crawler technology.

Secondly, extract file system and kernel from firmware with the extractor that FIRMADYNE developed based on the binwalk API. Due to the difference of firmware structure, the extractor cannot extract all firmware, such as QNAP's firmware, we use the specific firmware extracting program "extract_qnap_fw.sh" [14] invoking PC1 tool to achieve firmware extraction. After all the firmware is extracted, it is normalized and stored in the firmware database.

Finally, we simulate the NAS device based on FIRMADYNE. We define the kernel and libcnvram.so according to the file system extracted from the firmware, and hijack the NVRAM related operations, and then perform the system level simulation based on the QEMU system mode. We analyze the extracted kernel and find that the NAS firmware is mostly ARM and X86 architecture, while FIRMADYNE only supports ARM and MIPS kernel architecture, so we mainly choose the NAS device of ARM architecture for simulation.

Repeat the above steps to simulate the NAS devices of various manufacturers to form the Virtual NAS Set. The Virtual NAS Set has a strong scalability, and can continuously add various virtual NAS devices when the hardware resources are allowed.

Attack Recognition Center. Attack Recognition Center mainly combines attack chain to identify attacker's attack mode and method, including 3 modules: CVE Vulnerability Response Module, Resource Management Module, and Virtual Host Module. CVE Vulnerability Response Module contains a NAS vulnerability database, we get vulnerabilities information from CVE official website, and store vulnerabilities in

accordance with the mode of attack chain (① → ②→③ → ④) into the database. When an attacker attacks with a vulnerability in the NAS vulnerability database, the CVE Vulnerability Response Module compares the attack paths. If it is consistent with the way in which the vulnerability is exploited, it is considered to exploit the known vulnerability attack; if the path is inconsistent, it is considered to exploit the known vulnerability to expand the attack. Through resource access control, the Resource Management Module sets files and data accessed by administrators or specific users in the virtual NAS devices. If the controlled resources are accessed, it proves that the attacker has succeeded in invasion. If the attack path is: ① → ②→③, it is regarded as a device attack by password cracking; if the attack path is: ① → ③, an unknown attack is considered to have occurred. The Virtual Host Module mainly sets up the virtual host connected to the virtual NAS device. As a user of the NAS device, it monitors the attacker's penetration attacks through NAS. If the virtual host is attacked, it will be regarded as a successful penetration attack. The attack path is: ① → ② ③ → ④ or ① → ③→④. The Attack Recognition Center sends all the above attack information to the data center.

4 Experiments

4.1 Preparation

In order to detect the capture of attack behavior and the effect of system analysis, the NAS honeypot was built and connected to the Internet to attract attacks against NAS devices. Different attack behaviors are analyzed by capturing the attack behavior of the honeypot and capturing the system traffic. Select a router and connect to a host for traffic monitoring, because of that the NAS devices need to connect to the Internet through routing. And then connect the NAS honeypot to the network.

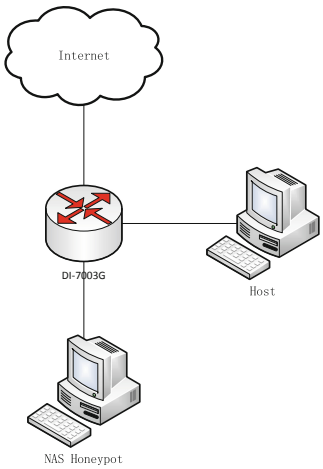

Fig. 3. The deployment of the experimental environment.

The deployment of the experimental environment is shown in Fig. 3. The equipment used is as follows:

(1) Two Windows10 hosts: one for NAS honeypot deployment and one for router traffic monitoring.

(2) One D-Link router, DI-7003G: used to connect the honeypot system to the Internet.

From 8:30 a.m. on March 24, 2020 to 8:30 a.m. on April 24, 2020, the system was deployed for one month. A total of 3053 NAS detection data were collected, among of which 2,247 successfully realized honeypot attacks.

4.2 Results and Analysis

In terms of the source of attack, the 3053 collected data came from 423 IP addresses in 41 countries and regions. Figure 4 shows the countries and regions with the source IP number in the top 10.

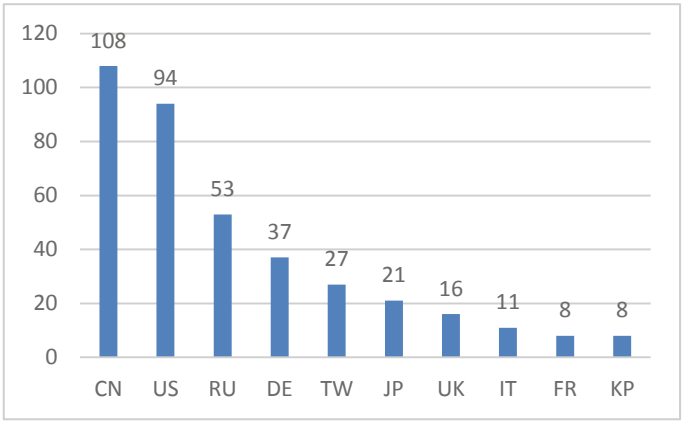

Fig. 4. Countries and regions with the source IP number in the top 10.

In terms of protocol utilization, the detection protocol based on SSH protocol is detected 1291 times, 1434 times based on HTTP protocol, 123 times based on FTP protocol, 16 times based on SMTP protocol, and 189 other probes, and the data statistics are shown in Fig. 5.

In terms of vulnerability utilization, according to the classification of attack types in NAS honeypot data center, 1587 of the 2247 successful attacks have been attacked by known vulnerabilities, 267 of which have expanded attacks, 655 have been attacked by device cracking, and 8 unknown attacks. In 1587 known attacks, gain privileges 183 times, bypass something 169 times, code execution 383 times, gain information 462 times, directory traversal 157 times, overflow 230 times, and data statistics are shown in Fig. 6.

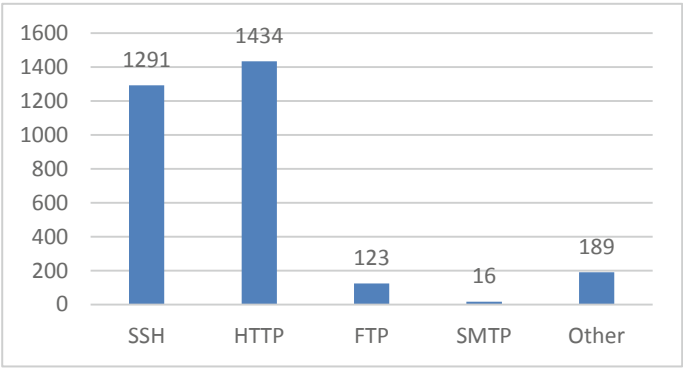

Fig. 5. Statistics of detection data protocol.

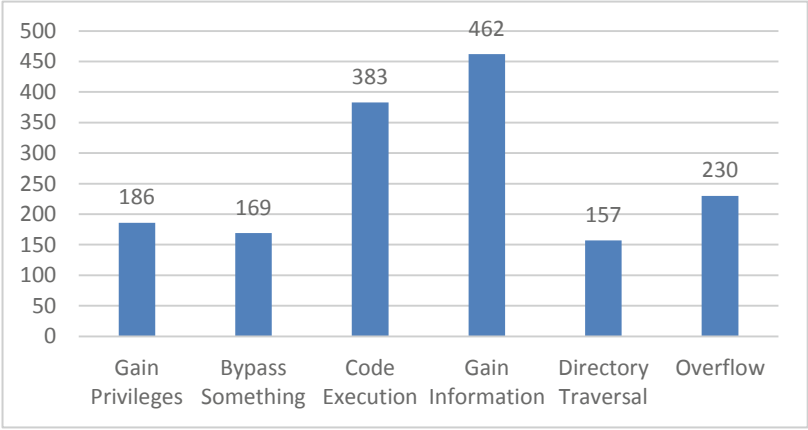

Fig. 6. Statistics of CVE vulnerability exploitation.

Honeypot system currently only provides NAS simulation of 3 vendors, including Synology, TerraMaster, and QNAP. The distribution of 2247 attacks in vendors: Synology 638 times, TerraMaster 529 times, and QNAP 1080 times, as shown in Fig. 7.

Select an example of QNAP attack which is successfully captured to prove the process of NAS honeypot catching, analyzing and recognizing the attack. The attack occurred at 13:00, April 15. The attacker detected the NAS device by sending the ICMP packet, the NAS honeypot system captured the detection data packet, used the QNAP response packet in the external fingerprint database to respond, and launched virtual QNAP NAS. The Virtual Response Layer forwarded attacker IP to the virtual QNAP NAS. The attacker did not engage in a password cracking attack. About 7 min later, the attacker exploited the vulnerability of CVE-2018-14746 to access the folder "import" that we had set up to view only with administrator privileges, and downloaded the 10 documents that were pre-created in the folder, through the way of command injection. Then, the attacker visited other folders without further attack actions, and the

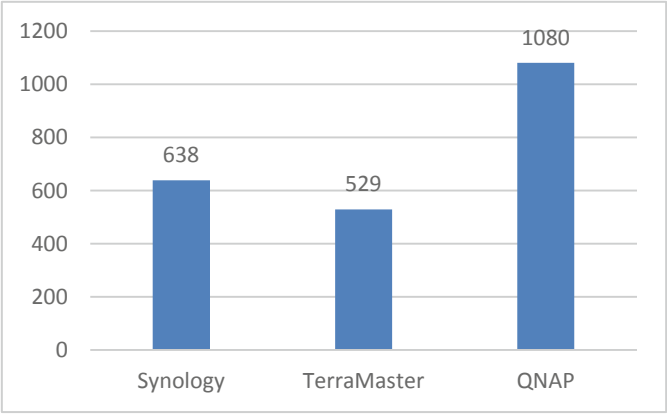

Fig. 7. Statistics of vendor attacks.

virtual NAS captured the above attacks. The NAS honeypot recorded the attack path of the attacker, ① → ③→④, which was consistent with the attack path in the CVE vulnerability database, and we judged that the attack behavior was using the known vulnerabilities to attack. The NAS honeypot maintained the virtual NAS to 18:00. No further attacks had been detected, and the honeypot had been automatically recovered to prepare for the next attack capture.

5 Conclusion

In this paper, NAS honeypot technology is studied based on threat model, and prototype system is implemented. The feasibility and flexibility and expansibility of the prototype system are verified by experiments. It can dynamically adjust the honeypot operation strategy, trap and analyze attacker attacks, and help to fully understand the security threats faced by NAS. Next, we will do further research on the technologies of NAS firmware virtualization for multi architecture, so that the honeypot can simulate more manufacturers and types of NAS devices to capture more unknown attacks.

Acknowledgment. This work was supported by the National Key Research and Development Program of China (2016YFB08011601).

References

1. Jiang, Z.: Network storage server advantage analysis. Modern economic information (03), 214 (2010)
2. Introduction to storage knowledge.https://wenku.baidu.com/view/efabd90e240c844768eaeea9.html. Accessed 20 May 2020
3. CVE-2018-18471. http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-18471. Accessed 16 June 2020

4. CVE-2016-10108.http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2016-10108. Accessed 16 June 2020
5. CVE-2020-9054. http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2020-9054. Accessed 16 June 2020
6. CVE-2018-13291. http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-13291. Accessed 16 June 2020
7. CVE-2018-13322.http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-13322. Accessed 16 June 2020
8. CVE-2018-14749. http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-14749. Accessed 16 June 2020
9. Moore, C.: Detecting ransomware with honeypot techniques. In: Cybersecurity and Cyberforensics Conference, Amman, Jordan, pp. 77–81 (2016). https://doi.org/10.1109/ccc.2016.14
10. Saud, Z., Islam, M.H.: Towards proactive detection of advanced persistent threat (APT) attacks using honeypots. In: 8th International Conference on Security of Information and Networks, Sochi, Russia, pp. 154–157 (2015). https://doi.org/10.1145/2799979.2800042
11. Anirudh, M., Thileeban, S. A., Nallathambi, D.J.: Use of honeypots for mitigating DoS attacks targeted on IoT networks. In: International Conference on Computer, Communication and Signal Processing, Chennai, India, pp. 1–4 (2017). https://doi.org/10.1109/icccsp.2017.7944057
12. Threat Modeling. https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff648644(v=pandp.10). Accessed June 2016
13. Chen, D.D., Egele, M., Woo, M., Brumley, D.: Towards fully automated dynamic analysis for embedded firmware. In: Proceedings of the 23rd Network and Distributed System Security Symposium (NDSS 2016). San Diego, CA: Internet Society, pp. 21–37 (2016)
14. https://github.com/max-boehm/qnap-utils