The Investigation and Development Status of Deep Packet Inspection

Xing ZHENG\textsuperscript{1,a,*}, Juan CAO\textsuperscript{1}, Hai-Jun GAO\textsuperscript{1}, Jie JIN\textsuperscript{2}

\textsuperscript{1}Key Laboratory of RF Circuits and Systems of Ministry of Education, Hangzhou Dianzi University, Hangzhou, 310018, China
\textsuperscript{2}School of Electronic Engineering, Dublin City University, Dublin 9, Ireland

\textsuperscript{a}zhengxing@hdu.edu.cn
\textsuperscript{*Corresponding author

Keywords: Deep Packet Inspection, Pattern Matching, Network security, Internet.

Abstract. The network security is an important issue in the Development of the Internet of Things (IoT). For detecting and defending the network attacks, the deep packet inspection is a promising technology that attracts the research and development attentions from the universities and industries. This paper overviews the investigations and development status of the deep packet inspection technology and summarizes different deep packet inspection techniques. Moreover, the paper also presents our investigations on the deep packet inspection and the related works.

Introduction

In recent years, the Internet of Things (IoT) develops very rapidly. It is estimated that \cite{1}, to 2020, the number of the access devices of the global IoT will increase to 7 billion. In the IoT, the core networks are still Internet. Therefore, the operations of the large number of the sensors are supported by the routing, switching and management devices. This raises a new problem: the network security problem in Internet will affect the operations of the sensor devices in the IoT. For example, the sensors in the smart home may have adverse impact on the people’s life when the attack occurs in the networks.

In the current Internet, the detections of the network attacks are mainly based on the upgrades of the virus databases and firewall settings of the end-user devices. However in the future, the core of the network security detection will move on the switching and routing devices. In the router end, the traditional packet detection is based on detecting the header information of the packets. This omits the network attacks hiding in the payload of the packets. For remedying the drawbacks of the traditional packet detection, the deep packet inspection have been proposed and obtained a lot of attentions from the academic society and industry circle. The deep packet inspection \cite{2} technology can realize: when the packet passes by the detection point, the header and the payload of the packet will be detected wholly. Therefore, the deep packet inspection can detect and defense the network attacks to the utmost extent.

The investigations of the deep packet inspection mainly includes: (1) packet classification which operates detection and matching for multi-fields of the packet header; (2) pattern matching which operates detection and matching for the payload of the packet. In packet classification, the techniques include \cite{3} for example, ternary content addressable memory (TCAM)-based packet classification, Bloom filter-based packet classification and decision tree-based packet classification. In pattern matching, the technique includes for example, fixed string-based pattern matching and regular expression-based pattern matching. Especially, the regular expression-based pattern matching includes deterministic finite automata (DFA) and nondeterministic finite automaton (NFA) \cite{4}.

Recently, NetFPGA \cite{5} which is an open source and reprogrammable hardware platform is widely investigated and used in the world. The advantages of using NetFPGA platform includes: (1) NetFPGA provides a fast and convenient way to implement a practical router and switch; (2) the NetFPGA can support modular feature, which simplifies the procedures of research and development;
(3) the architecture of the NetFPGA-based router is similar with the practical router. Therefore, it is feasible to use NetFPGA to investigate the deep packet inspection technology.

In this paper, the investigations and development status of deep packet inspection are presented in section 2, where different deep packet inspection techniques are described. Then, our investigations about the deep packet inspection and the related work are presented in section 3. Our research platform is also described in section 3. Finally, the paper is drawn in conclusions.

The Development of the Deep Packet Inspection

The existing deep packet inspection technology can be divided into two categories including packet classification and pattern matching. The packet classification aims at detecting and matching multi-fields of the packet header and the pattern matching aims at detecting and matching the data in the packet payload.

In the packet classification domain, most of the existing research work are based on the TCAM. This is due to the reason that the TCAM can perform the comparisons in parallel and report the matching result rapidly. However, the TCAM has the drawbacks of high power consumption and memory space. Moreover, the TCAM is not scalable with the clock rate and circuit area. The Bloom filter-based packet classification has low computation complexity of $O(1)$ and high memory efficiency [3]. But, the Bloom filter-based approaches also has some disadvantages, for example it does not support deterministic performance and is inefficient in handling with wildcard and prefix matching; it also needs extra modules to resolve the false positives, which limits the performance of the total system.

In recent years, the decision tree-based packet classifications attract more and more attentions from the researchers. The decision tree-based packet classifications use the geometric way to operate the packet classification. The decision tree-based algorithm cuts the rule space (corresponding to the packet header fields) into several smaller subspaces and it allows the matching process in a fast linear search way. Two classical decision tree-based packet classification algorithms are HiCuts and HyperCuts. HiCuts uses the local optimization decisions to choose which dimension to cut and how many cuts to make. The HyperCuts can cut the space on several fields per step, which can result shorter decision tree and support faster matching speed. Some works in recent years propose new packet classification techniques based on the decision tree. For example, [3] proposes a multi-pipeline architecture to enlarge the memory utilization while sustain high throughput.

In the packet payload detection, the traditional way is the fixed string matching. This matching scheme compares the packet content with sets of string, which is however limited by the match speed and scalability. The popular way of the payload matching is regular expression based pattern matching. It is an efficient detection way with improved expressiveness. The existing regular expression pattern matching are typically DFA [6] and NFA [7]. For a $n$-length regular expression, typical NFA has the operational complexity of $O(n^2)$ and the memory utilization of $O(n)$; typical DFA has the operational complexity of $O(1)$ and the memory utilization of $O(\Sigma^n)$ ($\Sigma$ is a set of strings).

Due to the reason that the detection and matching must be at high speed, DFA is widely used by more researchers. The memory space of DFA is determined by the product of the amount of states and the amount of transitions from each state. Therefore, the memory consumption is increased significantly when the size of DFA is increased. For addressing this drawbacks of DFA, $D^2$FA is proposed, in which several state transitions are replaced by a default state [8]. The main disadvantage of $D^2$FA is the increased matching time because the $D^2$FA has to traverse many states if it treats one character. For solving this problem, delta finite automata (named $\delta$FA) is proposed. In this scheme, most of the adjacent states are deleted as they share several common transitions and only the different states are kept.

The Experimental Setup and the Results

For investigating the deep packet inspection and the related work, we built an experimental testing environment, as shown in Fig.1. The experimental environment consists of a packet generator and a
router platform. One NetFPGA board is configured as a router. Another NetFPGA board is configured as a packet generator. To host the NetFPGA boards, two personal computers (PCs) with 2.93GHz Intel CPUs are used.

The packet generator acts as the traffic source and sink, with four ports connected to the router. The packet generator can create a packet stream and adjustable data rate by using the PCAP files. After the packet stream is forwarded by the router, the data rate can be calculated effectively at the corresponding destination port of the generator.

![Figure 1. The Experimental Setup.](image)

The deep packet inspection is being investigated in our labs. Here, a FPGA-based regular expression scheme is study and the simulation results are given in Fig. 2. In this figure, the deep packet inspector is successfully detecting the regular expression pattern of the ssh network protocol.

![Figure 2. The Simulation of the Deep Packet Inspection.](image)

For studying the deep packet inspection and other network security techniques, a high performance packet generator is urgently required to evaluate the performance of the network systems and devices. In our labs, a dynamic packet generator is implemented by using NetFPGA platform. In this scheme, the packet generator can dynamically change its throughput. Fig.3 Shows an example of the dynamic packet generator. The packet generator increases its speed almost linearly and can achieve 1Gb/s line rate atmost.

![Figure 3. The Increased Throughput of the Packet Generator.](image)

**Summary**

The deep packet inspection is one of the most promising technologies that can detect and prevent the network attacks efficiently. This paper overviews and describes a number of deep packet inspection technology. In addition, the paper also presents our experimental setups and some investigations of the deep packet inspection and the related works.
Acknowledgement
This work was supported by Zhejiang Provincial Qianjiang Talent Plan (QJD1402001), Zhejiang Provincial Natural Science Foundation of China (ZJNSF) (Grant Number: LQ13F010009), National Natural Science Foundation of China (Grant Number: 61331006), the National Natural Science Foundation of China (No.61372021), Science Foundation Ireland under the International Strategic Cooperation Award (Grant Number: SFI/12/ISCA/2496).

References