ISBN: 978-981-18-3959-7
2022 the 12th International Workshop on Computer Science and Engineering (WCSE 2022)
doi: 10.18178/wcse.2022.06.034

Research on Asynchronous Detection Method Based on Content Accelerator Card

Tang Haiyang+, Zhang Yingying and Jiang Jianjun

Shanghai DianJi University, Shanghai, China

Abstract. This paper analyzes the shortcomings of the existing content accelerator card technology[1],proposes an asynchronous detection method based on content accelerator card,and introduces the key technology of using content accelerator card to realize asynchronous detection in detail.The characteristics of the Netlogic's content accelerator card transform the content detection process,change the original synchronous detection mode into asynchronous detection mode,improve the utilization efficiency of the content accelerator card,greatly improve the product performance,and find a better solution for network security detection.

Keywords:accelerator card,asynchronous detection,network security.

1. Introduction

With the continuous development of network information security requirements,the traditional security products such as firewalls which act on layer 2 and layer 3 of TCP / IP protocol stack have been unable to meet people's needs.Now people are more concerned about what content is transmitted on the network and whether it will cause harm.Therefore, security products based on content detection such as IPS[2](Intrusion Protection System)、AV(Antivirus),which are gradually becoming the mainstream.However,one key issue that has limited the development of such products is performance.

The core of content detection is string pattern matching,which consumes CPU resources,and the CPU consumption increases linearly with the increase of pattern strings.[3]Therefore,only by breaking through the performance bottleneck of pattern matching and finding a better performance detection method,which makes the security products based on content detection get further development.

2. A Brief Introduction to Content Accelerator Card

Content accelerator card based on ASIC[4](Application Specific Integrated Circuit) and PCI-E[5] technology is invented to solve the above problems.Content accelerator card takes advantage of the characteristics of special hardware,which has the characteristics of strong processing power,processing speed does not decrease with the increase of pattern string and low cost.Content accelerator card connects to network security devices via the PCI-E interface,and it is also easy to extend existing network security devices.

The core of content detection is string pattern matching,which can use content accelerator card hardware to improve performance.However, content detection is a complicated process,which contains a lot of steps and string matching is just one of them,the rest of the work also need to deal with the CPU.A process in the CPU deals with a portion of the work, gives it to the content accelerator card for processing,and then returns it to the CPU again.The interaction between CPU and content accelerator card directly affects the efficiency of the whole working process and is also the key to improve the performance of content detection by using content accelerator card.

The CPU can interact with the content accelerator card in two ways.One is synchronous mode,that is,the CPU processes part of the work,when finished,the CPU hands it over to the content-accelerated card for

+ Corresponding author. Tel.: + 18251519323; fax: +38223108.
E-mail address: 1720545497@qq.com.

processing and then returns it to the CPU again.The whole processing is complete and continuous without interruption.This method has a small change to the detection process.However,due to the delay characteristics of content accelerator card,CPU needs to wait for a period of idle time and the efficiency is not high.The other is asynchronous mode,the whole detection process is divided into two parts,one half is responsible for the early processing and then gives it to the content accelerator card,the other half is responsible for taking out the results from the content accelerator card to continue the follow-up work,the two parts run independently without waiting.This method eliminates idle time of CPU waiting and improves detection efficiency and performance,but it also destroys the integrity of the detection process and is relatively complicated to implement.

The working process of the content accelerator card is generally as follows:

(1) Convert the pattern string to the internal presentation format of the accelerator card,which is loaded into the card by the accelerator card driver.If there are more than one pattern string,use different ID numbers to distinguish them.

(2) The process sends the matching string to the accelerator card and waits for the driver to return the matching result.

There are two ways to wait for the driver to return the result:one is synchronous mode,which is to send the string to the accelerator card,and the process blocks until the accelerator card returns the result.The other is asynchronous mode,where the string is sent to the accelerator card and the process continues to perform other operations,but the driver notifies the process to fetch the result when it comes out.

The drawback of the existing technology is due to the physical properties of the accelerator card,which has a large delay from sending the string to producing the result.Therefore,the synchronous mode of CPU utilization is lower,and asynchronous mode can give play to the maximum characteristics of the accelerator card.

3. Asynchronous Detection Method Based on Content Accelerator Card

3.1. Asynchronous Content Detection Techniques

Figure 1 shows a schematic diagram of realizing asynchronous detection by using content accelerator card.In the figure,the data packet is received from the network card,passed through the network module,and then delivered to the user process module.The user process module calls the accelerator card module to finish pattern matching,and finally it is returned to the network module and sent out from the network card.

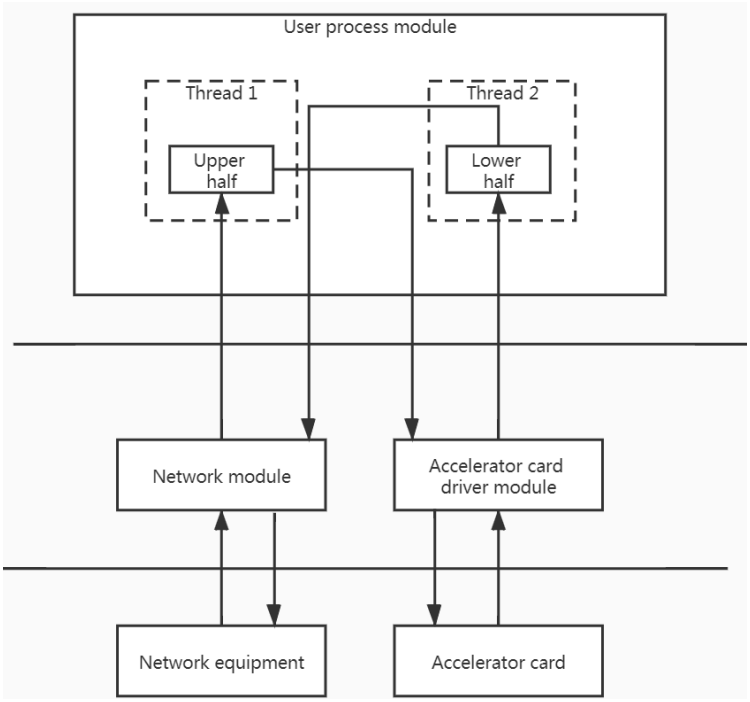

Fig. 1: Schematic diagram of asynchronous detection

The work of the user process module includes four steps:receiving data packet,preprocessing data packet,pattern matching and processing results.These four steps are divided into two parts.The upper half part includes receiving data packet and preprocessing data packet.The lower half part includes pattern matching and processing results.

Thread 1 handles the upper part and Thread 2 handles the lower part. The two parts together complete a integrated data packet content detection process.However,the work of the upper and lower parts is carried out simultaneously without interfering and waiting for each other. After the preprocessing of the upper part,the interface function provided by the content accelerator card is called to deliver the data packet to the content accelerator card,and then the next data packet is immediately processed without waiting for the result.

The lower part calls the interface function provided by the content accelerator card to get the matching result.If it gets the matching result,it will continue to process according to the matching result,such as checking the header of the message,calling other third-party plug-ins, taking blocking actions or recording logs,etc.If the result is not obtained, the match is unfinished and no processing is done at this time.

The point here is that content accelerator card often have multiple detection channels and can accommodate multiple messages to submit matches at the same time.The results obtained in the lower part may be multiple,that is,the matching work of multiple messages is finished and the results are generated at the same time.At this time,the lower part needs to process continuously,and the results can be obtained after multiple results are processed.

When a datagram is submitted to the content accelerator card for pattern matching,a datagram flow ID is returned,which is the unique identity of the datagram.When the matching result is obtained,the identification is also taken,which is mainly used to match a data packet with a matching result,especially when multiple matching results are produced at the same time,to uniquely identify the corresponding relationship.

The work of multiple matching channels in the content accelerator card is done simultaneously,which greatly improves the performance compared to the CPU detection.However,the matching speed is different for different data packets.Even if the data packets to be matched are submitted one by one according to the sequence of channels,it is possible that the data packets submitted later will be matched first and generate the result,that is,it does not conform to the "first-in,first-out" rule.This is mainly caused by the different length of message and the different complexity of message content.

3.2. Asynchronous Detection Algorithm

Figure 2 shows the flow chart of asynchronous detection which use the content accelerator card.First of all, the user process module starts and performs initialization.Converts the pattern string to an accelerator card internal format and is loaded into the accelerator card by the accelerator card driver during initialization.

Content accelerator cards of different technologies have their own specifications,which are primarily a subset of the regular expressions that can be supported.For example,Netlogic's accelerator card is a single chip structure,PCI-E interface,with a large capacity of pattern matching ability and can support the complete PCRE[6].

The compiler software provided by the accelerator card is used to compile the rules in the rule base during initialization,and convert them into pattern features that can be recognized by the content accelerator card.Then the interface function provided by the accelerator card is called to load the compiled pattern features onto the content accelerator card.If the features are more complex and numerous,the loading process will be longer.Loading is usually a one-time event and is not reloaded unless the rule base is updated.

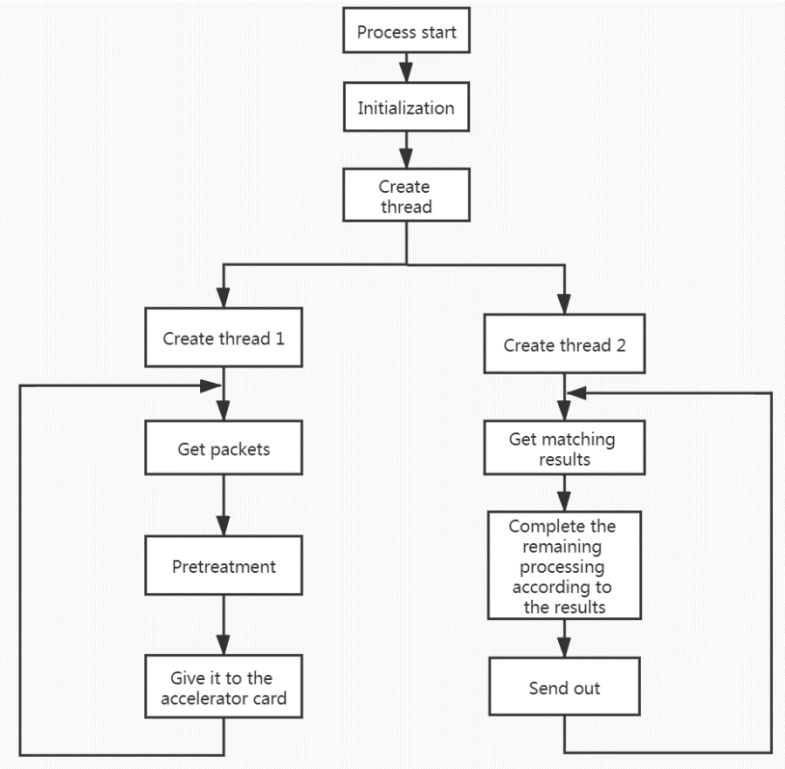

Fig. 2: Flow chart of asynchronous detection

The process module starts two threads:Thread 1 and Thread 2,and they work on different tasks.

The thread 1 process is as follows:

(1) Get the data packets from the network module.

(2) Perform DDOS[7] detection and its attack is usually in the form of a flood,that is,a large number of requests are made to the target host by different sources at the same time,making the target host resource overloaded and unable to handle normal access requests.DDOS needs to be detected before pattern matching.One reason is that DDOS takes up a large amount of bandwidth and poses a threat to the security device itself.It needs to be detected as early as possible at the beginning of the program and blocked in time.Another reason is that sending a DDOS message to a content accelerator card for pattern matching would greatly reduce the efficiency of the content accelerator card,and make it do a lot of useless work due to DDOS attacks cannot be detected by matching.

(3) It is processed at the network level,and the data packets without data content need not be detected and be forwarded directly.Management messages submitted to the device itself also do not need to be detected and be submitted directly.There are also some routing exchange information messages,Layer 2 broadcast messages and so on.The principle of processing is that the messages,which do not need to be detected should not be sent to the subsequent detection program as far as possible in order to reduce the pressure on the detection program.

(4) For data packet preprocessing,protocol analysis is the first step.IP segmented messages need to be reorganized,while TCP messages need to record connection information in order to detect data content in a stream.HTTP,POP3,SMTP,FTP,and other commonly used application layer protocols need to be parsed and disassemble the header structure that needs to be detected.

(5) Generates private information associated with the packet as follows: struct ips.IPS is added to private data that on the connection and one for each connection.

```
struct ips {
    u32 session;    /* Handle to the connection structure */
    u32 policy_id[2];      /* Policy_id [0] where the policy ID is matched,*/
    u8 policy_dir; /* Direction of policy matching,1:c->s;2:s->c; 3: c<->s */
    u8 ids:1;  /* IDS Mode*/
```

```
u8 fastpath:1;  /* Take the fast path and don't do any testing */
u8 anyany:1;
u8 established:1;    /* The TCP connection has been marked */
u8 nat:1;  /*    The connection did the NAT */
u8 check_nat:1;
u8 handle_cpu;      /* All packets for this connection must be sent to the CPU,which determined by the packet distribution policy*/
u32 natport[2];
u32 nataddr[2];      /* The address to replace when NAT */
u16 packets[2];      /* Record the number of packets detected in Direction 1 and Direction 2, respectively */
};
```

(6) The interface function provided by the content accelerator card driver is called to send the detected message to the content accelerator card for pattern matching.The stream ID parameter is provided at the time of the call so that the content accelerator card matches against the data stream rather than a single message.That is,after matching a message,a segment of data at the tail of the message (the length can be configured) is saved and matched with the second header of the same data stream next time.In this case,the attack features can be found when they span between two messages.

(7) Go back to step (1) and continue processing the next data packet.

The thread 2 process is as follows:

(1) Get the matching result of the data packet from the accelerator card driver.When the accelerator card completes matching to a data packet,it generates an interrupt.The driver hands the result to thread 2,which is waiting for the result,and wakes up thread 2 to work.

(2) Thread 2 finds the corresponding private information of the data packet and continues the unfinished processing.

(3) Thread 2 returns the message to the network module for sending or discarding according to the result of processing.

(4) Go back to step (1) and continue to deal with the next result.

4. Conclusion

Using the content accelerator card of NetLogic,and taking advantage of the characteristics of the accelerated card,the content detection process is transformed by the method.The original synchronous detection mode is changed into asynchronous detection mode.The receiving preprocessing of the datagram is separated from the matching result and the subsequent processing of it.Two parts work asynchronously,which improves the utilization efficiency of the content accelerator card and greatly improves the product performance.

There is another obvious advantage to using content accelerator cards for asynchronous detection.Even with the addition of a few detection rules,performance can remain almost the same.And with the increase of the number of detection rules,the detection performance of traditional CPU synchronous detection method will be greatly reduced.Even when the number of rules increases to a certain level,the device is not performing well enough to continue processing data packets.With the emergence of network attacks,the detection rule base also needs to be constantly updated and the number of rules is constantly increasing.With the increasing demand for content security detection technology,the application of this research technology can maximize the detection performance of products and reduce the cost of products.

5. Acknowledgements

Fund project:The next generation Internet technology innovation project of Cernet NGII20170505; Funded by Shanghai Science and Technology Commission, project: Research on generic key technologies for smart family services, No14511108001

6. Reference

[1] Chen Ruining.Research on Network Application Acceleration Based on QAT Virtualization [D].Shanghai Jiao Tong University,2019.

[2] Zhang Sheng,Shi Ronghua,Zhao Jue.Research and application of radial node link graph in IPS log analysis [J].Journal of Central South University (Science and Technology),2017,48(07):1774-1781.

[3] Gu Zhaojun,Guo Jingxuan.Internal Threat Detection Method Based on Role Abnormal Behavior Mining.Computer Engineering and Design,2020,41(10):2740-2746.

[4] PU Tian-lei.Design and Key Technology Research of ASIC Chip for Front End Readout of Gas Detector [D].University of Chinese Academy of Sciences (Institute of Modern Physics, CAS),2020.

[5] Li Ming.Research on Key Technologies of High Performance Password Card Based on PCI-E [D].Xidian University,2018.

[6] Xiang Longgang,Ge Huiling.A Trajectory Movement Mode Analysis Method Based on Directional Quadrant Mapping [J].Geomatics and Information Science of Wuhan University,2020,45(04):495-503.

[7] Wang Wentao,Li Shumei,Tang Jie,Lv Weilong.DDoS Attack Detection Method Combining Probability Graph Model and DNN [J/OL].Computer Engineering and Applications :1-17[2021-01-14].