2011 International Conference on Information Management and Engineering (ICIME 2011)
IPCSIT vol. 52 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V52.28

The Research and Implementation of UTM-HA

Chen Zhang [+], Bin Hou, Gang Liu

School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China

Abstract. Since emerged in 2004, Unified Threat Management (UTM) has been brought into wide use to enhance network security protection. In practical use, UTM can not only be treated as gateway to control the connection of inner net and outer net, but also as safety devices providing many services and functions including access control, intrusion prevention , anti-virus and so on. Therefore, the much higher request is bring up for the high-availability of UTM. To address this issue, this paper presents UTM-HA (UTM High Availability), a solution to improve the high-availability of UTM. The solution has been implemented and its performance was evaluated.

Keywords: high-availability; network; security; UTM;

1. Introduction

With fast development of the Internet, network security becomes an important issue, as large numbers of security threats exits in current Internet, such as virus, worms, Trojans and malicious attacks. As a primary network gateway defense solution, Unified Threat Management (UTM) emerged in 2004, and has been used widely. According to IDC, the official definition of UTM is: “Products that include multiple security features integrated into one box. To be included in this category, an appliance must be able to perform network firewall, network intrusion detection and prevention and need not be used concurrently, but the functions must exist inherently in the appliance.” [1-2]

The biggest value with UTM platforms is simplicity and lower price given its “all-in-one” footprint [3]. As a gateway device, UTM can block network threats at gateway before they have the opportunity to enter your network or attack individual desktop PCs or mail servers. Nevertheless, “all in one” footprint reduces UTM’s ability to withstand risks and increases the possibility of the single point of failure in the whole system [4]. As a gateway device, any single point mistakes in UTM system will not only result in the invalidation of the safety measure, but also influence the normal processing of the entire network. In order to solve the above problems, we concentrate our attention on the high-availability of UTM.

The official definition of system availability, especially electronic systems, is: “a ratio of the expected value of the uptime of a system to the aggregate of the expected values of up and down time.”[5] According to this definition, system availability is just a function of time without consideration of performance. Based on the general availability definition of system, we consider that availability of UTM should include two aspects, one is called time availability which evaluates the running persistence of UTM, another is called capacity availability assessing the ability that UTM provides the stated function and performance, and propose the solution to improve the high-availability of UTM named UTM-HA. Our solution includes two main parts: 1) Analysis and Judgment of Trouble. 2) Trouble-shooting.

In summary, we have made the following contributions:

- Alternative solutions to recurring different troubles are proposed.

[+] Corresponding author. Tel.: + (15210833672).
E-mail address: (zhang1988chen@126.com).

- Combination of fault tolerant technique and Duplex Hot-Backup System is presented.
- System implementation is introduced and demonstrates that the solution can be implemented with the qualities of high-availability, sustainability and synchrony.

The remainder of the paper is structured as follows. Section 2 presents an overview of related work. Next, Section 3 discusses the design of control mechanism. We present the implementation in Section 4, followed by the experiment and evaluation in Section 5. We briefly discuss our on-going work in Section 6.

2. Related Work

To meet the growing demands for high-availability of UTM, UTM manufacturers have proposed a broad variety of solutions, which can be divided into two types: Bypass mode and Duplex Hot-Backup mode [6]. These solutions are introduced in this section.

2.1 Bypass Mode

With the help of embedded Switch Circuit, UTM will make two-port all-pass networks and pass all traffic unconditionally when abnormal circumstances and exception conditions occur in UTM system such as power failure, reboot, crashes and so on.

Although this mode is easy to realize and many Industrial PC (IPC) have implemented hardware bypass, it is not a perfect choice for UTM which is treated as network safety equipment including firewall function. The default function of firewall is to block all traffic unless users configure it to pass, while Bypass mode will pass all, which is in contradiction with firewall and not able to fulfill people’s security requirements. In other words, Bypass couldn’t reach the goal of high-availability.

2.2 Duplex Hot-Backup Mode

Using two or more UTM equipments connected by heart-throb line to establish synchronization sessions and monitor states of the other party, this mode automatically switches between the host and the backup machine or takes over the defective appliance to guarantee the services available when UTM device fails.

This mode can still fulfill people’s security requirements when abnormal circumstances and exception conditions occur in UTM, but this mode is difficult accomplished and it’s a waste to use the backup UTM taking over the host UTM if there is only a small problem with the host one. We should feel out the situation first before we take any definite steps, that is to say, we only adopt Duplex Hot-Backup mode when it is necessary.

3. Design of UTM-HA

3.1 Design Goals of UTM-HA

We did some analysis on the desired requirements that UTM-CM must supply. We now list them below.

- Sustainability. Once server exception occurs, services will be stopped for a period of time, which is hard to bear for some enterprise applications. So it is highly important to provide continuous services, even providing 365 × 24 services[7]. Therefore, one of the vital requirements of UTM-HA is sustainability that UTM can still provide basic security functions when there is trouble and trouble-shooting.
- Synchrony. If the configuration of the backup UTM is different from the host one, it may not work normally when switching between the host and the backup one. What’s more, the threat signatures and the lib of intrusion rule should also be synchronously updated.
- Alert function. UTM-HA should send out a warning to administrator so as to repair the machine in good time as it automatically choose the most effective solutions to ensure the system operating properly if exception conditions occur.

3.2 Hardware Structure of UTM-HA

UTM-HA is a harmonious combination of fault tolerant technique and Duplex Hot-Backup System. As shown in Figure 1, UTM-HA adopts Duplex Hot-Backup System as whole framework [8], consisting of two special designed UTMs, which have fault tolerant technique.

According to the location of server data, there are two kinds of method to realize Duplex Hot-Backup System, common shared memory mode, with the same storage device for the host and backup server, and data synchronization based on software mode. The latter mode is more suitable for UTM, for the following reasons: UTM only need to store some proxy software configuration information and a little space will be enough; common shared memory mode increases the possibility of the single point of failure in the whole system. For the above-mentioned reasons UTM-HA adopts data synchronization based on software mode, which will be explained in detail in Software Structure of UTM-HA.

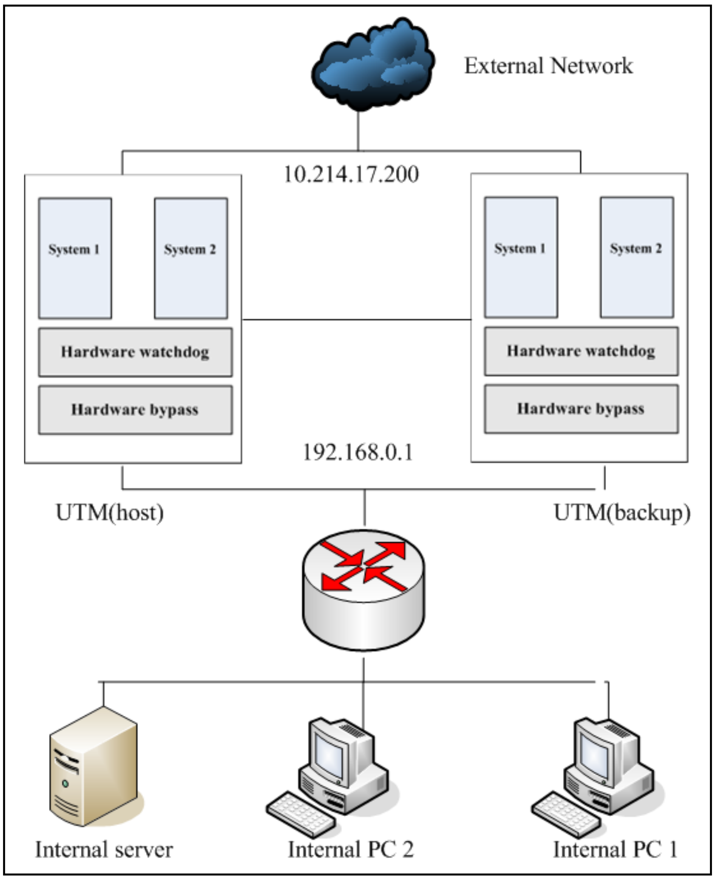

Fig.1: Hardware structure of UTM-HA

There are two special designed UTMs in UTM-HA, as shown in Figure 1, each UTM containing hardware watchdog and hardware bypass. Hardware watchdog, monitoring the hardware state of UTM, will trigger hardware bypass to pass all traffic through this UTM when UTM cannot work normally, which ensures the normal processing of the entire network during switching between the host and the backup UTM.

3.3 Software Structure of UTM-HA

In order to achieve high-availability, we not only adopt Duplex Hot-Backup System as whole framework but also designs special software structure of UTM. We now present an overview of the software structure of UTM (see Figure 2). It is described on a high-level overview in this section. We will discuss the details and implementation of several components in the following sections.

We divide UTM into four parts:

- GUI. It is human-to-machine part and receive user configurations about UTM.[9]
- Security Center. The center contains system1, as host system, system 2, as backup system, and configuration storage. System 1 and system 2 contain the same security modules, such as anti-virus, firewall, IPS, and so on and they have the common configuration storage to store rules and user

configurations. When system 1 malfunctions, system 2 will read configurations from the common configuration storage to maintain unity with system 1 and take over system 1 to ensure UTM operate normally.

- Inside Monitor. Software watchdog which monitors the states of security modules, such as anti-virus, firewall, IPS and so on will trigger software bypass to restart the failure security module when there is only one security module out of work, or to pass all traffic and start system 2 when there is more than one failure module.
- Outside Monitor. It consists of heart-throb monitor and data synchrony. The host UTM and the backup one are connected by heart-throb line and heart-throb monitor monitors the other UTM's state. At the same time, the host UTM data synchrony module will transfer data to the backup one when some changes take place to keep the same configuration between the host one and the backup one.

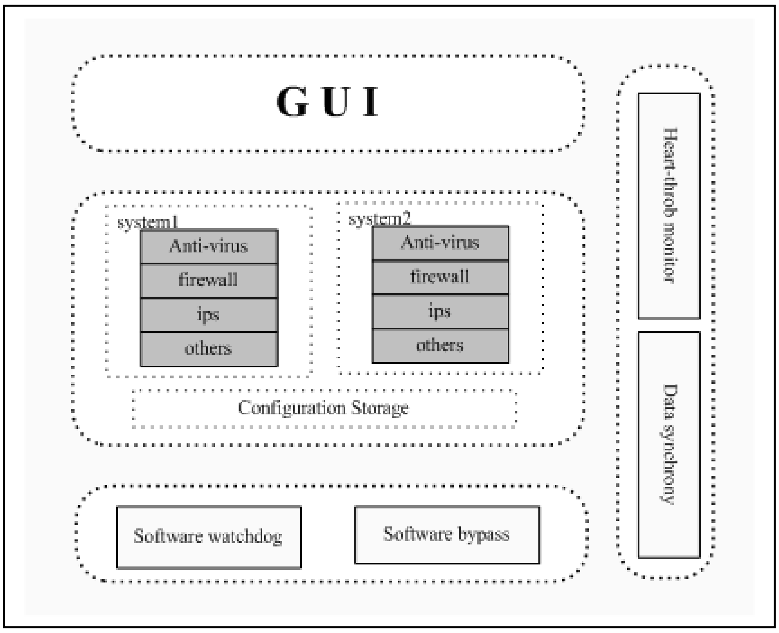

Fig.2: Software structure of UTM

UTM-HA are made up of two UTMs connected by heart-throb line and the software structure is the same as UTM. It will alert when there is exception and alert function module is not included in Figure 2.

4. Implementation Details

4.1 Watchdog & Bypass

- Software watchdog and bypass. The chief reason for implementation of software watchdog and bypass is to enhance the repairing ability of single UTM. From Figure 3, we can clearly see the total process of the implementation of software watchdog and bypass. We use process 1 to N symbolizing security modules and use Monitor process to monitor watchdog and check registering information including process id and system time delivered by process 1 to N. Firstly, Monitor process starts automatically whenever you start system1 and continues to run as long as you are using system 1.Then produce 1 to N will also automatically start and each produce has a counter. Each produce sends registering information every 1s and counter automatically adds one when sending registering information succeed. At the same time, monitor process makes decision according to the number of counters which are greater then zero.
- Hardware watchdog and bypass. Traditionally, when watchdog takes effect, system 1 will reset. But in UTM-HA, we use watchdog writing low level to GPIO (General Purpose Input Output) by which bypass is controlled. Watchdog triggers bypass not reset and the network still work during the backup UTM takes over the host one.

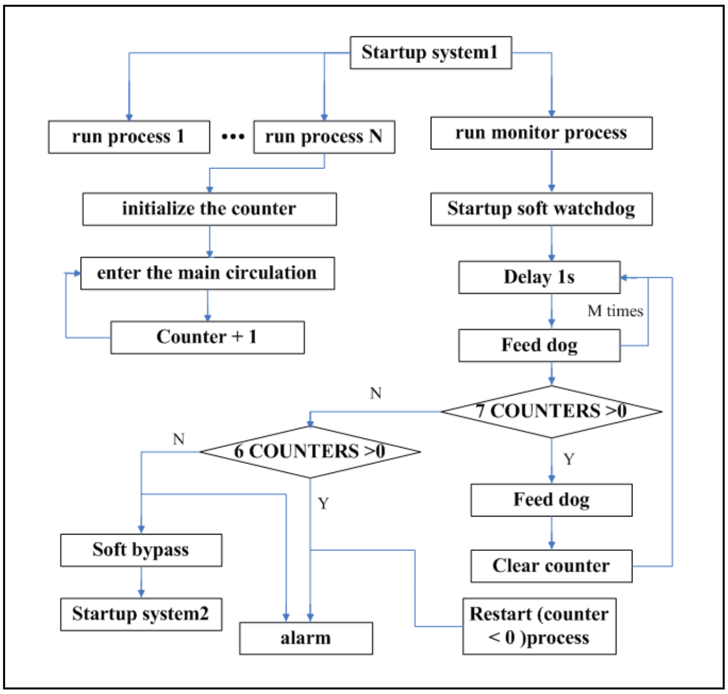

Fig.3: Software watchdog model implementation

4.2 Duplex Hot-Backup Mode

As mentioned in Section Ⅲ, UTM-HA adopts data synchronization based on software mode as Duplex Hot-Backup Mode. Implementation of the mode will be introduced in detail in the following section.

- Data synchronization. Only when threat signatures, the lib of intrusion rule or configuration of the host UTM have changed, is variation transferred to the backup UTM. The backup one will receive incomplete data when exception occurs during the transmission, which may result in inconformity. In view of the above-mentioned facts, we apply the transaction concept, that is to say, the transaction will be called off when transmission fails to keep synchrony.

- Data transfer safety. Some confidential data are transferred between the host and backup UTM through network. In UTM-HA, SSH2, which adopts symmetric-key algorithms, such as AES, DES, 3DES, Blowfish, RC4 and so on, is used to protect the security of data transfer. Besides, DSA and DH algorithms instead of RSA are used to exchange symmetrical key and HMAC instead of CRC is used to assure data consistency.

- Test for the host UTM activity. The most effectual way to test whether the host UTM is out of work is heart-throb mechanism. The heart-bolt program of the host UTM regularly sends the host UTM’s state while the heart-bolt program of the backup UTM real-timely monitors the host UTM’s state sent by the host one. The backup one will consider the host one out of work and take over the host one when the backup one doesn’t receive the host UTM’ s state within a predetermined time period.

- When for taking over the host UTM. The host UTM is so busy that it may not send heart-bolt signal within a long time period, which called suspended animation. If the backup one takes over the host one the instant that the backup one doesn’t receive the heart-bolt signal from the host one, it may result in frequently transition between the host and the backup UTM and interrupting the service. So we apply a buffer window for the backup one. When the backup one doesn’t receive the heart-bolt signal within a predetermined time period, it will send an urgent command to the host one, and the host one must give the command priority. After that, if the backup one has not yet receives the heart-bolt signal, it takes over the host one.

4.3 Exception Handling System of UTM-HA

We group exceptions occurred in UTM-HA into four principal categories and present the relevant resolve measures.

Case1: One security module is out of work. Alert and restart the security module.

Case2: More than one security module of the host system is out of work while the host system is active and the backup is standby. Alert and startup the backup system and software bypass.

Case3: More than one security module of the backup system is out of work while the host system is broken. Alert and startup the other UTM and hardware bypass.

Case4: Power off or hardware faults of UTM. Alert and startup the other UTM and hardware bypass.

5. Experiment and Evaluatuation

5.1 Experiment

To objectively evaluate the function and performance of UTM-HA, a series of tests based on shaking hands every second and 9600 bps through heart-throb line have been carried out. We simulate all kinds of failure to test weather UTM-HA can automatically handle exceptions and get the time cost in handling these exceptions.

Case1: Man-made one security module out of work.

Case2: More than one security module of the host system is out of work while the host system is active and the backup is standby.

Case3: More than one security module of the backup system is out of work while the host system is broken.

Case4: Power off or hardware faults.

In the above cases, UTM-HA alarms automatically and handles the problems properly: In case1, Restart the failure security module; in case2, Startup the backup system and software bypass; in case3, Startup the other UTM and hardware bypass; in case4, Startup the other UTM and hardware bypass. The time cost in switching over system 1 and system 2 is about 2 seconds and over the host UTM and the backup UTM about 3 seconds. During the switchover process, the network can still work normally.

5.2 Evaluation

The goals of designing UTM-HA mentioned in Section Ⅲ have been accomplished, which can be seen from the above experiment. Next, we evaluate UTM-HA from different angles.

Two popular methods of evaluating availability are the number of nine and DPM (Defect Per Million) as shown in Table Ⅰ. The number of nine, measuring availability in percentage, is often used to evaluate the availability of host of big computer [10]. In comparison, DPM is more suitable to measure the availability of network, including many types of equipment.

Based on the above analysis, we adopt DPM as our evaluation method. But we just take three months not million hours for sample to make the statistics of the defect. There are three categories of defects: Restart security module, happened 9 times, Startup the backup system 7 times, and Startup the backup UTM 2 times. Figure 4 shows the distribution of the exception: Restart security module comprises 50 per cent, Startup the backup system 39 per cent, and Startup the backup UTM 11 per cent.

From the above we can see that many failures can be fixed by the host UTM and need not activate the backup UTM. That is to say, UTM-HA first tries to fix the failure by the host one and doesn't stop sending heart-bolt signal to activate the backup one until the host one can not fix the failure. By doing that, UTM-HA can not only improve the high-availability but also reduce cost.

TABLE I. METHODS OF EVALUATING AVAILABILITY

Availability	*DPM*	*Interrupting time per year*
99.000%	10000	3days,15hours,36 minutes
99.900%	1000	8hours,46minutes
99.990%	100	53minutes
99.999%	10	5minutes

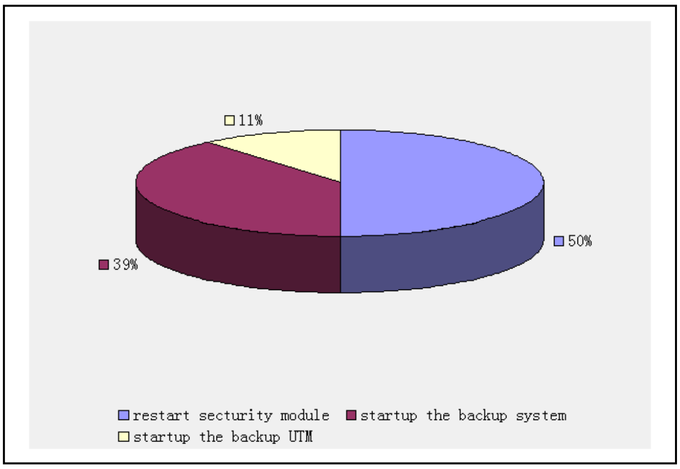

Fig.4: The distribution of exception of UTM-HA

6. Conclusions and Future Work

To reach high-availability Unified Threat Management, fault tolerant technique and Duplex Hot-Backup System should be combined rather than simply adopted separately. In this paper, we presented UTM-HA to improve the high-availability based on that two techniques and provided alternative solutions to recurring different troubles.

In the future, we would like to implement complicated QoS on UTM. We would also like to study UTM based on collaborative work model.

7. References

[1] http://www.idc.com/.

[2] http://www.venustech.com.cn/.

[3] http://www.freescale.com/.

[4] Ming-jun Ling, “Research on UTM Technology”, 1009-8054(2008) 09-0064-02.

[5] She-sheng Gao, Ling-xia Zhang, “Availability Theory and Enginneering Application”, Academic Press, 2002.

[6] http://safe.zol.com.cn/109/1098042.html.

[7] T. Chou, “Beyond fault tolerance”, In IEEE Computer 30(4):31-36, 1997.

[8] P. Neira, Laurent Lef`evre, R.M. Gasca, “High Availability support for the design of stateful networking equipments”, Proceedings of the First International Conference on Availability, Reliability and Security (ARES’06).

[9] Ying Zhang, Fachao Deng, Zhen Chen, “UTM-CM: A Practical Control Mechanism Solution for UTM System”, 2010 International Conference on Communications and Mobile Computing.

[10] http://www.cisco.com/web/CN/products/products_netsol/switches/pdf/core_sl_04.pdf.