2010 International Conference on Broadband, Wireless Computing, Communication and Applications

Darknet Monitoring on Real-Operated Networks

Seiichiro Mizoguchi
Kyushu University
744, Motooka, Nishi-ku,
Fukuoka, Japan
mizoguchi@
itslab.csce.kyushu-u.ac.jp

Yoshiro Fukushima
Kyushu University
744, Motooka, Nishi-ku,
Fukuoka, Japan
yfukushima@
itslab.csce.kyushu-u.ac.jp

Yoshiaki Kasahara
Kyushu University
6-10-1, Hakozaki, Higashi-ku,
Fukuoka, Japan
kasahara@
nc.kyushu-u.ac.jp

Yoshiaki Hori
Kyushu University
744, Motooka, Nishi-ku,
Fukuoka, Japan
hori@inf.kyushu-u.ac.jp

Kouichi Sakurai
Kyushu University
744, Motooka, Nishi-ku,
Fukuoka, Japan
sakurai@inf.csce.kyushu-u.ac.jp

Abstract

Darknet monitoring is an effective method to analyze malicious activities on networks including the Internet. Since there is no legitimate host on darknets, traffic sent to such a space is considered to be malicious. There are two major issues for darknet monitoring: how to prepare unused address space and how to configure network sensors deployed on the network. Preparation of monitoring addresses is difficult, and it have not been obvious yet what an appropriate configuration is. To solve the first issue, we proposed a method for network monitoring by exploiting unused IP addresses on segments managed by DHCP server, where is a real-operated network. By assigning these addresses, we can easily obtain IP addresses for monitoring and enable network monitoring on production network. Furthermore, we conducted real darknet monitoring experiments and clarified what kind of information could be obtained. We deployed several types of sensors on real-operated network and captured darknet traffic. After analyzing the traffic, we compared the data between each sensor. We found that there were dramatic differences between the data collected by each sensor and our proposed method was useful for real network monitoring.

1. Introduction

With the development of computer networks including the Internet, the number of adversaries on the Internet is also increasing. They attack computers connected to the network and try to compromise them. They develop their own tools for attacking, such as viruses, worms, and bots. These malicious software are called "Malware," and new types of malware are constantly appearing. For example, a bot is one of the most considerable issue in the world. A host infected by them can be controlled by adversaries and conduct network scanning attacks has the self-update function and always revise their binary code in order to avoid anti-virus software detecting them.

In this situation, there is malicious traffic on the Internet generated by these malware. Network monitoring is one approach to see them. There are unused IP addresses on the Internet called a darknet, where several packets from the Internet can be seen. Since there is no legitimate host on the darknet, these packets are considered to be malicious. Several researchers have tried to capture these packets by monitoring the network traffic and analyze them[4, 5]. They obtain information about network attacks from captured traffic. That information will be used for intrusion detection, grabbing the situation of the network.

In network monitoring, some special devices called network sensors are used. There are two kinds of network sensors, a darknet sensor and a honeypot. A darknet sensor has a function that only capture packets sent to it. There is no interaction between sensors and attackers. The information obtained from one darknet sensor is a little bit small, but it is cheap and easy to be deployed. A honeypot can communicate with attackers. Since a honeypot has more functionality than darknet sensors, it can obtain detailed information about malicious activities. However, it is much more difficult to deploy and maintain honeypots.

978-0-7695-4236-2/10 $26.00 © 2010 IEEE
DOI 10.1109/BWCCA.2010.82

There are two major issues to be considered in network monitoring. One is the preparation of unused address spaces for monitoring, and the other is the configuration of network sensors to be deployed. Generally, we have to prepare unused address blocks for network monitoring, however, it is difficult to prepare such a network in organizational networks or universities. Furthermore, it is necessary to decide proper parameters of network sensors in order to get valuable information, but so far, it is not clear what kinds of information can be obtained from darknet monitoring.

In this paper, we proposed a method for network monitoring which exploits unused IP addresses on real-operated network: a production network. Using unused IP addresses on segment managed by DHCP (Dynamic Host Configuration Protocol) service, we could easily obtain addresses for monitoring and deploy network sensors. Furthermore, in order to solve the issue about configuration of network sensors, we revealed the differences of information between several types of network monitoring system. We conducted real network monitoring, putting several types of network sensors on our campus network, and analyzed the data collected by the sensors. We found that there were dramatic differences between the data collected by each sensor and our proposed method was valuable for real network monitoring.

2. Network Monitoring

2.1. Two Major Issue

We have to consider two major issues: preparation of IP addresses for monitoring, and configuration of network sensors. Generally, unused address blocks are prepared for network monitoring. If there is a real host on the network which is monitored, sensors may receive noise of traffic related to that host. By using a segment only for network monitoring, pure darknet traffic can be obtained. Moreover, if we deploy honeypots, it is important to isolate monitoring segments from real operating network. If a honeypot is compromised and exploited as a stepping stone, it may attack local production computers. However, it is difficult to prepare such a network in organizational network or universities.

The other issue is configuration of network sensors. Darknet traffic observed at each sensor may be dramatically different and the differences are due to the environment where these sensors are. Especially, the place where the sensors are deployed is an important factor that affects to capability of obtaining valuable information. It is said that, attackers decide targets by using some algorithms, for example, random choice of IP addresses or sequential one. If attackers conduct network scan with completely random method, we have to put as much network sensors as we can in order to catch these attacks. However if the network scans are sequential one, we do not have to deploy so many sensors to monitor the attacks. While, even if we have already known the attackers' scanning method, we do not know how many sensors we should deploy on the network. Furthermore, the characteristics of network attacks may differ due to the situation of network. For instance, if there are many servers on network segment, the segment may be attacked with much higher rate than the segment which has production computers. In the related works, the closeness of network sensors and live networks seems to be one factor that affects to the visibility of network threats. Attackers may change their scanning methods to ensure the effectiveness of their attacks.

2.2. Related Works

Ruoming Pang et al show the characteristics of traffic sent to unused address spaces[8]. They defined such traffic as "Background Radiation" and they collect them from four unused network in the Internet. Then they identified the components of background radiation by protocol, application, and specific exploit, and extract the patterns of background radiation and activities. Evan Cooke et al reported the result of blackhole placement[6]. They have deployed a large size of network monitoring system called Internet motion sensor (IMS) [5]. There are four main contributions. First, they deployed ten distributed blackhole sensors at major service providers, large enterprises, and academic networks. Second they identified sensor placement as an important factor in understanding and generalizing darknet measurements. Third, they provided strong empirical results showing differences between traffic observed on diverse distributed blackhole sensors, and the last they proposed definition and application of sensor properties to explain differences in traffic measurement on darknets. Artail et al proposed a hybrid honeypot framework which improves intrusion detection systems in organizational networks[3]. In order to monitor as much unused IP addresses as they can, they proposed deploying a lot of honeyd in production network. Honeyd is a low-interaction virtual honeypot and a lot of honeypot daemons are executed on one machine[1]. However, in order to implement their system, they have to prepare a lot of static IP addresses. They prepared such IP addresses on production computers and assigned them to honeyds. In their work, they conducted network monitoring and analyzed the data collected by their monitoring system. They reported that one honeyd emulating a domain name server tends to be attacked more than the others, and other honeyds emulating HTTP is also attacked frequently.

3. Monitoring on Production Network

3.1. Basic Idea

As we mentioned in previous section, unused spaces for monitoring should be prepared in order to get pure darknet traffic. However, we believe that monitoring on live network is also important. Attackers will intentionally conduct various attacks on live network, so the information obtained from live network can be much valuable. If we deploy darknet sensors or honeypots on live network and monitor only the traffic toward to those sensors, we can eliminate noises to some extent and get pure darknet information. Thus in this work, we focus on unused IP addresses located between live hosts.

On live network, a DHCP service is often used to manage IP addresses for production computers. If a computer is connected to the network, DHCP server will automatically assign an IP address to it. Several IP addresses are reserved for DHCP service. However, there are potentially many unused addresses. To obtain addresses for monitoring, we make use of these addresses.

Next considerable issue is configurations of network sensors. In order to get information that you want, we have to choose appropriate configuration, however such a configuration is not obvious. We have to find proper settings that match the environment where you are going to monitor.

To solve this issue, we conducted several experiments on our campus network, which is real-operated and many production computers exist. By deploying network sensors in real network and capturing darknet traffic, we attempt to clear what kind of darknet traffic can be obtained and how to configure the network sensors. We will analyze obtained data and compare the result with previous works. Then we will investigate appropriate configuration for network monitoring.

3.2. Experimental Setups

We prepared three monitoring segments in our campus network. One of them is a /24 size of network that are not used. We captured packets from the Internet to this unused network. Since there is no legitimate host on that network, an edge router of our campus replies to darknet packets with host unreachable message. We monitored darknet traffic at this edge router. Second and third one are the network sensors on production network where we routinely use. We deployed 20 darknet sensors and 20 honeypot sensors on that network and assigned unused IP addresses managed by DHCP service.

We used a honeyd, a low-interaction virtual honeypots[1] for darknet sensors and honeypots. We can easily configure interaction level of each honeypot. Darknet sensors have no interaction, which means they do not reply any packets from the Internet. On the other hand, honeypots have a little interaction, that is, they reply only one time to the packets. In detail, they send TCP/RST packets for TCP/SYN, ICMP port unreachable packets for UDP, and ICMP echo reply for ICMP echo request. We captured all packets destined to these sensors with tcpdump[2].

There is one limitation. Our campus network has a firewall and packets sent to vulnerable port such as TCP 23, 135 or 445 are filtered at the edge router. We can not capture these packets on production network.

Table 1. The total number of packet for each protocol and the average (packets/day)

Total (packets)	/24 unused N/W	Darknet	Honeypot
TCP	8375972	9138524	110036
UDP	506340	24421	22811
ICMP	159805	10383	10351
Average (packets/day)	/24 unused N/W	Darknet	Honeypot
TCP	90065	1257776	15145
UDP	5445	3362	3140
ICMP	1719	1430	1425

4. Experimental Result

We conducted our experiments for three months, 2009/10/10 – 2010/01/10. We will show the result in this section.

4.1. The total number of packet

At first, we focus on the total number of packets observed by each monitoring system. Table 1 shows the detail of the number and the average of number of packets, and Table 3 shows destination port numbers targeted most and its packet rate. The average is normalized by the size of /24 so different size of sensors can be compared. The number of packets is 256/20 times magnified for darknet sensors and honeypots.

Darknet sensors received much more TCP packets than /24 network and honeypots. The majority of these packets is the one toward TCP port 1433. This result seems to be due to the interaction level of the sensors. Honeypots replied TCP/RST packets to senders of those packets thus we can assume that the senders might gave up attacking those IP addresses because no service is provided in the IP addresses, while darknet sensors never replied to the senders thus the senders continued to compromise them. To confirm our assumption, we examined the average of the number of pack-

ets sent to a destination address (honeypot or darknet sensor) by a source address (attacker). If an attacker gives up attacking a host by receiving a TCP/RST packet, the number of the packets sent to the host by the attacker could be small. We calculated the average numbers of those values among 20 IP addresses (honeypots and darknets sensors) using 3 traffic data captured on Oct. 20, 2009, Nov. 1, 2009, and Jan. 5, 2010. We show the results in Table 2. From Table 2, we can see that the average numbers of honeypots are smaller than those of darknets.

Next, we focused on one source IP address and analyzed how many packets have been sent to each sensor by one source host. As a result, there are several source host that change their behavior of sending packets according to whether the sensor is a darknet sensor or a honeypot. For example, one source host (Type (A)) sent only one packet to each honeypot, while it sent dozens of packets to each darknet sensor. In this case, the attacker might give up attacking these honeypots anymore. However, there is another host (Type (B)) that had different behavior from Type (A). It sent much more packets to honeypots than darknet sensors. In this case, the host might give up sending packets to darknet sensors because they did not reply to the scan packet (These results are shown in Appendix A). However, from Table.2, the number of source hosts that behave like Type (A) is more than Type (B), thus it can be concluded that attackers truly give up attacking honeypots if the honeypots reply TCP/RST packets to the attackers. For this, the number of the packets captured on darknet sensors would be more than that of honeypots.

/24 unused network received packets toward TCP 445 and 135 so the average number of TCP packets is higher than honeypots. On the other hand, the average numbers of UDP and ICMP packets seem to be almost the same between three sensors.

Table 2. The average of the number of packets sent to a host by an attacker

	Honeypot	Darknet
10/20/2009	5.530	10.149
11/1/2009	5.331	33.490
1/5/2010	6.149	9.601

Since there are a lot of TCP packets toward port 1433, we focused on the number of these packets for each day. Packets toward TCP port 1433 is corresponding to a MS SQL server vulnerability[7]. Figure.1 shows graphs that represent the number of TCP 1433 packets observed by each sensor. In this figure, darknet sensors have constantly received much more packets than the others. From this analysis, the data captured between production computers can be dramatically different from the one observed by other sensors.

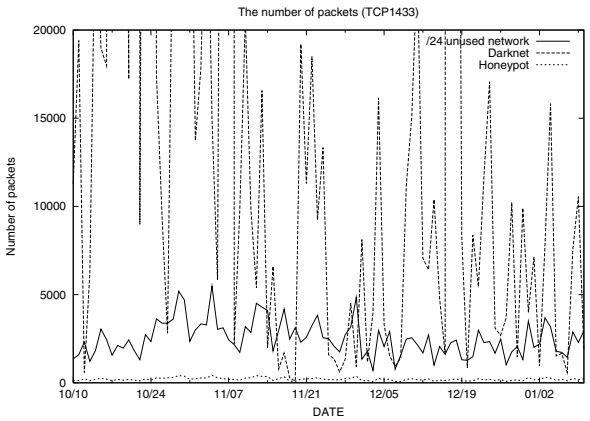

Figure 1. The number of packets focus on TCP port 1433. Darknet Sensors have received much more TCP 1433 packets than the others.

Table 3. Port numbers targeted most and Packet rate per sensor

/24 unused N/W		Darknet		Honeypot	
port	Rate (packets/sensor)	port	Rate	port	Rate
445	25471	1433	260022	1433	947
1433	912	80	244	22	727
2967	608	22	123	2967	593
135	352	2967	119	80	248
23	342	8080	95	1521	156

4.2. The number of packet for each sensor

Next, we focus on the number of packets for each sensor. Cooke et al mentione that the beginning of address block tends to recieve more packets than the tail of the block[6]. At our /24 unused network (Figure 2), a similar result was provided. Moreover, the shape of graph is remarkable. The sensors which are first half of block (.0 to .127) recieved much more packets than the rest. Especially, among the first half area of block, the beginning of the area received the most packets (.0 to .12). The reason seems to come from scanning method of attackers. Since it is not efficient to scan whole network, attackers may tend to attack the first part of address block where servers or other systems are often placed.

We conducted the same analysis to sensors on production network, and we found that the result of honeypot sensors resemble the result of /24 unused network. On the other

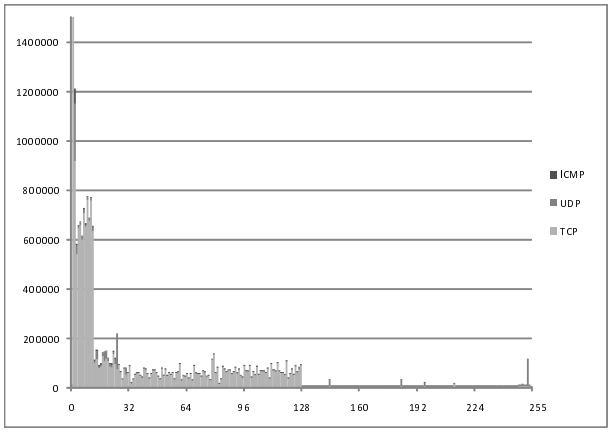

Figure 2. The number of packets for each destination address (/24)

hand, at darknet sensors, the number of packets and the place where they are deployed seems to be independent.

4.3. The number of unique source address

Next, we focus on the number of unique source address that are observed at each sensor. This number is valuable for detecting security incident on the Internet. For example, a botnet, which is a cluster of bots, conducts attacks simultaneously, so if it starts attacking, the number of source address will suddenly increase.

Table 4 shows the average of that number. When we focus on the source address that send UDP and ICMP packets, the numbers of each sensor are the same value, however the number of host sending TCP packets to /24 unused network is very large. The result might occur because of packet filtering at the edge router. The source addresses sending packets toward TCP port 445 or 135 has a great variety. On the other hand, since the numbers of source address that sent UDP and ICMP packets are the same for each sensor, they have almost the same ability of collecting darknet traffic if there is no packet filtering feature.

Darknet sensors and honeypot sensors could constantly capture about 200 source addresses per day. So there may be no relation between interaction level of sensors and the number of source address captured by them. However, for darknet sensor, they captured more source addresses that throw TCP packets than ones which throw UDP packets. This reason may be due to their interaction level. We have to make obvious about this result.

Figure.3 and Figure.4 show the number of unique source addresses observed in /24 unused network. In Figure.3, the number of host sending TCP packets to the network is so high. Most of them targeted TCP port 445. Figure.5 and

Table 4. The average number of unique source address.

	TCP	UDP	ICMP
/24 unused N/W	31700	1083	46
Darknet (normalized by /24)	1525	1006	92
Honeypot (normalized by /24)	1232	976	89

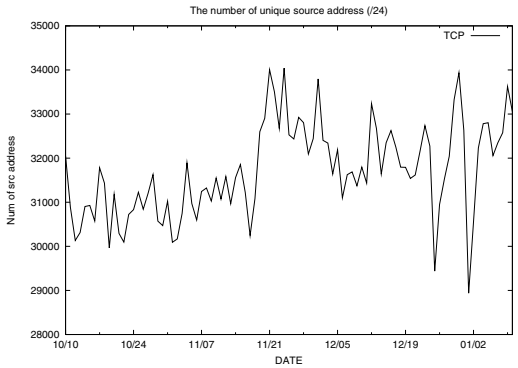

Figure 3. The number of unique source addresses observed in /24 unused network (TCP packets)

Figure.6 show those of darknet sensors and honeypots. The shapes of the two figures have similar characteristics, however, it seems that the number of source addresses monitored by darknet sensor is slightly more than that of honeypots. We need to more detail.

4.4. Discussion

The results are dramatically different between each sensor. The most remarkable one is the place where the sensors were deployed. The beggining of the network tends to be attacked more than the rest. We can propose one idea for network monitoring that network sensors should be placed on the first half of network in order to receive more packets from more source IP addresses.

This time, our experimental result was affected by the packet filtering. We did not eliminate the effect of filtering, so we have to apply elimination of these features. Another issue is localization of our monitoring system. We deployed network sensors on very samll part of our campus network. To generalize the result, we have to deploy another network sensor in different places. One more issue is interaction level of honeypot sensors. In this time, our honeypot can only reply TCP/RST which means there are no productive service, so we have to raise interaction level or provide real network service such as HTTP or FTP or SQL servers.

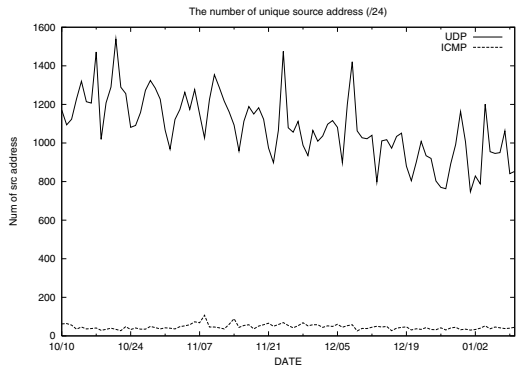

Figure 4. The number of unique source addresses observed in /24 unused network (UDP and ICMP packets)

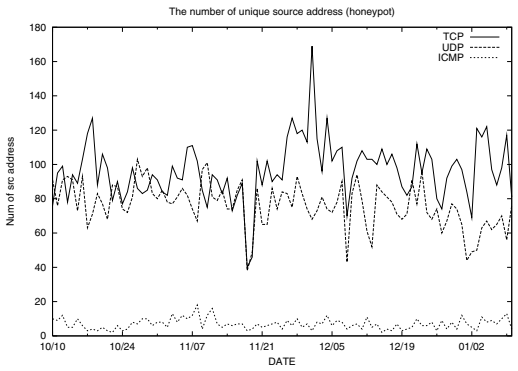

Figure 6. The number of unique source addresses observed by honeypots

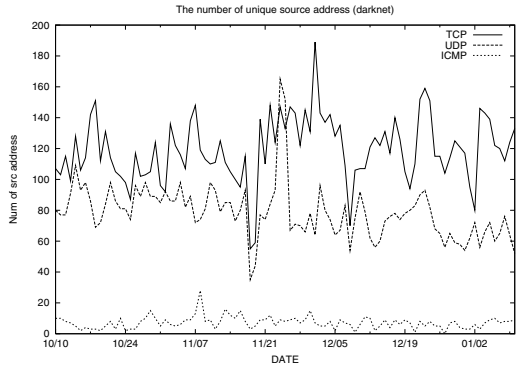

Figure 5. The number of unique source addresses observed by darknet sensors

5. Conclusion

Network observation technique is valid plainly when grasping the status of networks. However, there are two major issues: preparation of unused address spaces for monitoring, and determining appropriate configuration of network sensors which are deployed. We tried to solve these issues. First, we proposed a network monitoring method which exploits unused IP addresses on production network, where IP addresses are managed by DHCP server. Using unused IP addresses on segment managed by DHCP service, we need not prepare unused address spaces for network monitoring.

Second, in order to find appropriate configuration of sensors, we conducted real-operated network monitoring. We setup network sensors with several configurations and monitored darknet traffic on production network. We analyzed the data obtained by each sensor, compared the result of monitoring and evaluated our proposed method. We found that there were dramatic differences between the data collected by each sensor.

For future works, we have to do further consideration of deployments of network sensors. We expand our network monitoring system to whole campus network and try to conduct larger network monitoring. In this case, we have to think about collaboration of information obtained by distributed monitoring systems, and data reduction of collected by a lot of sensors.

Acknowledgment

This research was partly supported by a contract with the National Institute of Information and Communications Technology, entitled “Research and Development for Widespread High-speed Incident Analysis.”

References

[1] Developments of the Honeyd Virtual Honeypot, February 2010. `http://www.honeyd.org/index.php`.

[2] tcpdump/libpcap, May 2010. `http://www.tcpdump.org/`.

[3] H. Artail, H. Safa, M. Sraj, I. Kuwatly, and Z. Al-Masri. A hybrid honeypot framework for improving intrusion detection systems in protecting organizational networks. *Computers & Security*, 25:274–288, 2006.

[4] M. Bailey, E. Cooke, F. Jahanian, A. Myrick, and S. Sinha. Practical darknet measurement. In *40th Annual Conference on Information Sciences and Systems (CISS)*, pages 1496–1501, Princeton, NJ, 2006.

[5] M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson. The internet motion sensor: A distributed blackhole monitoring system. In *Proceedings of the 12th ISOC Symposium on Network and Distributed Systems Security (SNDSS)*, pages 167–179, 2005.

[6] E. Cooke, M. Bailey, Z. Mao, D. Watson, F. Jahanian, and D. McPherson. Toward understanding distributed blackhole placement. In *Proceedings of the 2004 ACM workshop on Rapid malcode*, pages 54–64, NY, USA, 2004.

[7] Z. Li, A. Goyal, and Y. Chen. Honeynet-based botnet scan traffic analysis. *Botnet Detection*, pages 25–44, 2008.

[8] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson. Characteristics of internet background radiation. In *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pages 27–40, Taormina, Sicily, Italy, 2004.

Appendix

A The number of packets sent to each sensor by one source host

Table.5 shows the number of packets which were sent from one source address to each sensor. ID 1 to 20 are darknet sensors and 21 to 40 are honeypots. Host1 to Host4 belong to Type (A). Type (A) host send one packet to honeypots and much more packets to each darknet sensor. This result might occur because honeypots return TCP/RST packets so those hosts gave up sending scanning packets anymore. Host5 is another type of host that sends more packets to each honeypot than to each darknet sensor (Type (B)). This host might begin another scanning attack to honeypots when it knows the existence of some interesting hosts. Host6 is another one which sends packets constantly regardless of darknet sensors or honeypots. We calculated the average number of packets for each sensor and its variance. If the variance is small, these sensors receive almost the same number of packets from the source host. For example, because each honeypot received one packet from Host1, its variance was 0. For Host4, since one darknet sensor (ID2) received much less packets than the others, the average was around 500 but its variance became very large. However, most of these variances are small values, so attackers may take the same actions to the same type of sensors.

Table 5. The number of packets sent to each sensor

	Darknet Sensors (ID 1 to 20)								Honeypots (ID 21 to 40)							
Source Host	1	2	3	···	19	20	Avg.	Var.	21	22	23	···	39	40	Avg.	Var.
Host1	8	8	9	···	8	8	8.05	0.045	1	1	1	···	1	1	1	0
Host2	29	29	29	···	34	34	32.6	4.948	1	1	1	···	1	1	1	0
Host3	5	5	5	···	5	5	5	0	1	1	1	···	1	1	1	0
Host4	558	158	522	···	560	538	522	7226	1	1	1	···	1	1	1	0
Host5	1	1	1	···	1	1	1	0	34	34	34	···	34	34	34	0
Host6	3	3	3	···	3	3	3	0	3	3	3	···	3	3	3	0
:				:								:				
:				:								:				