Network Analysis Report

1. Executive Summary

In a detailed examination of the network traffic data, this report aims at determining whether a single system in the observed environment was infected by a malware. As the investigation was being conducted, a single internal computer has an unusual behaviour in comparison to the behaviour of other systems on the network. The given workstation was seen to be in a state of constant communication with outside Internet addresses at the pattern, which can be described as characteristic of remote control mechanisms and automated and not the activity of an ordinary user. The rate and amount of these outgoing connections are a strong indication that the system is likely infected by botnet malware. Although no direct evidence exists to suggest that the vulnerable data of the traffic was exposed because of the presence of the given aberrant behavior, the observed one convincingly shows that the system was likely violated within the timeframe that was being analyzed, which requires urgent containment and remediation of the situation in an active environment.

2. Dataset Information

- File Name : botnet-capture-20110816-sogou.pcap
- File Size : 18 MB
- File Format : PCAP
- Encapsulation : Ethernet
- Snapshot Length : 65535 bytes

Capture Time Details

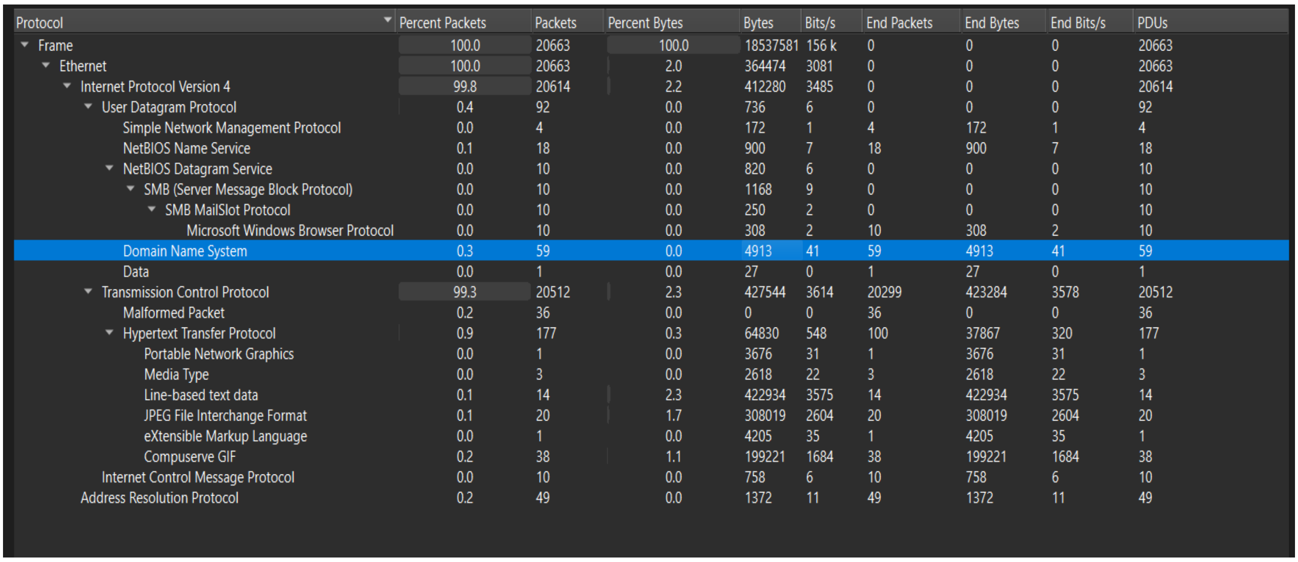
- First Packet : 2011-08-16 06:56:24
- Last Packet : 2011-08-16 07:12:10
- Capture Duration : 15 minutes 46 seconds (946.188 seconds)

Traffic Statistics

- Total Packets : 20,663
- Total Bytes : 18,537,581 bytes (18.5 MB)
- Average Packets/sec : 21.8 pps
- Average Packet Size : 897 bytes
- Average Throughput : 156 kbps

The dataset has a short acquisition time, of only 15 minutes and 46 seconds, and an average throughput of about 21.8 packets/sec. Average packet size is 897 bytes which suggest it maybe an application-layer transactions not a normal scanning operation or SYN flood. This suggests that the traffic level is relatively low, which likely indicates controlled bot activity or normal web exchanges rather than a large-scale volumetric attack.

3. Protocol Hierarchy



Protocol	Percent Packets	Packets	Percent Bytes	Bytes	Bits/s	End Packets	End Bytes	End Bits/s	PDUs
Frame	100.0	20663	100.0	18537581	156 k	0	0	0	20663
Ethernet	100.0	20663	2.0	364474	3081	0	0	0	20663
Internet Protocol Version 4	99.8	20614	2.2	412280	3485	0	0	0	20614
User Datagram Protocol	0.4	92	0.0	736	6	0	0	0	92
Simple Network Management Protocol	0.0	4	0.0	172	1	4	172	1	4
NetBIOS Name Service	0.1	18	0.0	900	7	18	900	7	18
NetBIOS Datagram Service	0.0	10	0.0	820	6	0	0	0	10
SMB (Server Message Block Protocol)	0.0	10	0.0	1168	9	0	0	0	10
SMB MailSlot Protocol	0.0	10	0.0	250	2	0	0	0	10
Microsoft Windows Browser Protocol	0.0	10	0.0	308	2	10	308	2	10
Domain Name System	0.3	59	0.0	4913	41	59	4913	41	59
Data	0.0	1	0.0	27	0	1	27	0	1
Transmission Control Protocol	99.3	20512	2.3	427544	3614	20299	423284	3578	20512
Malformed Packet	0.2	36	0.0	0	0	36	0	0	36
Hypertext Transfer Protocol	0.9	177	0.3	64830	548	100	37867	320	177
Portable Network Graphics	0.0	1	0.0	3676	31	1	3676	31	1
Media Type	0.0	3	0.0	2618	22	3	2618	22	3
Line-based text data	0.1	14	2.3	422934	3575	14	422934	3575	14
JPEG File Interchange Format	0.1	20	1.7	308019	2604	20	308019	2604	20
eXtensible Markup Language	0.0	1	0.0	4205	35	1	4205	35	1
Compuserve GIF	0.2	38	1.1	199221	1684	38	199221	1684	38
Internet Control Message Protocol	0.0	10	0.0	758	6	10	758	6	10
Address Resolution Protocol	0.2	49	0.0	1372	11	49	1372	11	49

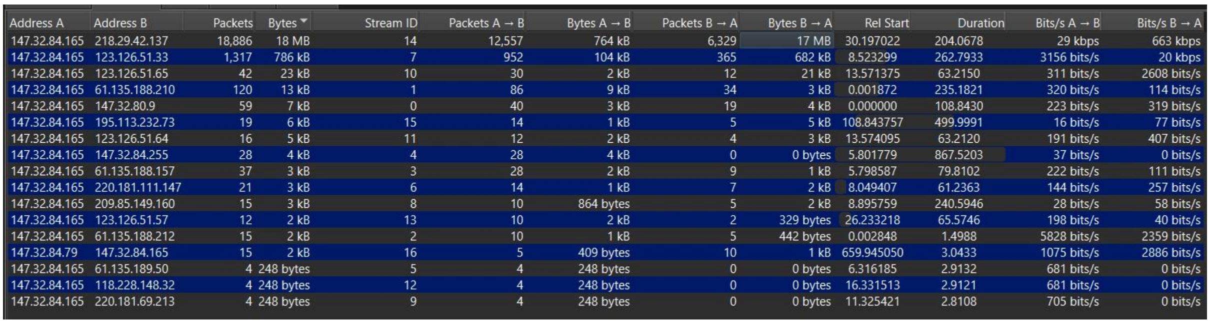
Figure 1

Key Protocol Distribution

Protocol	Percent Packets	Notes
TCP	99.3%	This indicates application layer communication most likely web browsing or HTTP based communication
UDP	0.4%	Minimum usage, rules out DNS tunnelling or UDP-heavy attack
HTTP	0.9%	There's small packet count but carries significant bytes shows presence of web content like image and text
DNS	0.3%	DNS activity is very low so may not be aggressive DGA or DNS beaconing
ARP	0.2%	Normal local network protocol activity

The traffic is mostly TCP based, which means that the majority of traffic is in the application layer. The small DNS usage rate suggest that host machines are not resolving new domains very often, thus the possibility of domain-based command-and-control activity is very low in this dataset. Although the volume of traffic on the HTTP protocol is very low in terms of the number of packets sent, it accounts disproportionately to the total data volume because of the presence of images, web pages, and text-rich content. The trend indicates that the web browsing or web-based communication is the dominant traffic. There is also no major TLS traffic, which is in line with a 2011-type of data where encrypted C2 or HTTPS-intensive communication was less common. Only a few bad packets (36, about 0.2%) are observed and they might be a result of capture artifacts or a bad traffic, which should be investigated further.

4. Top Talkers (IPv4 Conversations)



Address A	Address B	Packets	Bytes	Stream ID	Packets A → B	Bytes A → B	Packets B → A	Bytes B → A	Rel Start	Duration	Bits/s A → B	Bits/s B → A
147.32.84.165	218.29.42.137	18,886	18 MB	14	12,557	764 kB	6,329	17 MB	30.197022	204.0678	29 kbps	663 kbps
147.32.84.165	123.126.51.33	1,317	786 kB	7	952	104 kB	365	682 kB	8.523299	262.7933	3156 bits/s	20 kbps
147.32.84.165	123.126.51.65	42	23 kB	10	30	2 kB	12	21 kB	13.571375	63.2150	311 bits/s	2608 bits/s
147.32.84.165	61.135.188.210	120	13 kB	1	86	9 kB	34	3 kB	0.001872	235.1821	320 bits/s	114 bits/s
147.32.84.165	147.32.80.9	59	7 kB	0	40	3 kB	19	4 kB	0.000000	108.8430	223 bits/s	319 bits/s
147.32.84.165	195.113.232.73	19	6 kB	15	14	1 kB	5	5 kB	108.843757	499.9991	16 bits/s	77 bits/s
147.32.84.165	123.126.51.64	16	5 kB	11	12	2 kB	4	3 kB	13.574095	63.2120	191 bits/s	407 bits/s
147.32.84.165	147.32.84.255	28	4 kB	4	28	4 kB	0	0 bytes	5.801779	867.5203	37 bits/s	0 bits/s
147.32.84.165	61.135.188.157	37	3 kB	3	28	2 kB	9	1 kB	5.798587	79.8102	222 bits/s	111 bits/s
147.32.84.165	220.181.111.147	21	3 kB	6	14	1 kB	7	2 kB	8.049407	61.2363	144 bits/s	257 bits/s
147.32.84.165	209.85.149.160	15	3 kB	8	10	864 bytes	5	2 kB	8.895759	240.5946	28 bits/s	58 bits/s
147.32.84.165	123.126.51.57	12	2 kB	13	10	2 kB	2	329 bytes	26.233218	65.5746	198 bits/s	40 bits/s
147.32.84.165	61.135.188.212	15	2 kB	2	10	1 kB	5	442 bytes	0.002848	1.4988	5828 bits/s	2359 bits/s
147.32.84.79	147.32.84.165	15	2 kB	16	5	409 bytes	10	1 kB	659.945050	3.0433	1075 bits/s	2886 bits/s
147.32.84.165	61.135.189.50	4	248 bytes	5	4	248 bytes	0	0 bytes	6.316185	2.9132	681 bits/s	0 bits/s
147.32.84.165	118.228.148.32	4	248 bytes	12	4	248 bytes	0	0 bytes	16.331513	2.9121	681 bits/s	0 bits/s
147.32.84.165	220.181.69.213	4	248 bytes	9	4	248 bytes	0	0 bytes	11.325421	2.8108	705 bits/s	0 bits/s

Figure 2

Top Internal Hosts by Outbound Traffic

Internal IP	External IP	Packets	Bytes	Notes
147.32.84.165	218.29.42.137	18,886	18 MB	Dominant outbound traffic; potential automated communication or C2 activity.
147.32.84.165	123.126.51.33	1,317	786 kB	Moderate communication; secondary external endpoint.
147.32.84.165	123.126.51.65	42	23 kB	Low-volume connection; likely occasional traffic.

Observations :

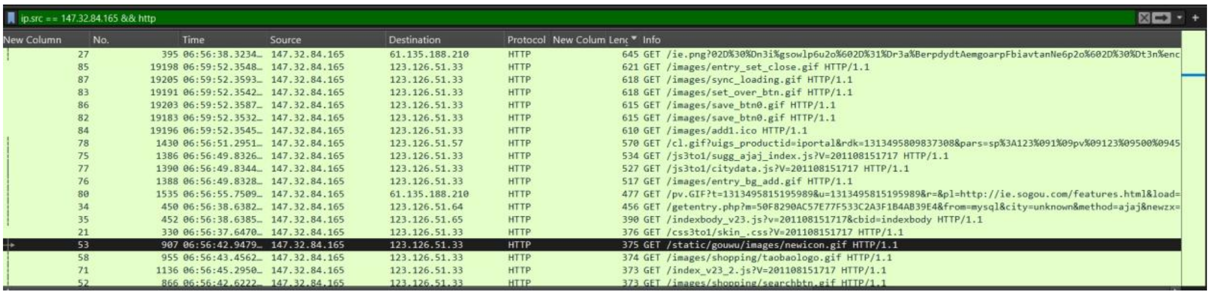
1. Single dominant host 147.32.84.165(internal IP): The single dominant host that is proposed in the capture is the single host with almost all the traffic in it, so it can be considered the primary candidate to the further behavioral analysis.

2. Several external endpoints: In the dataset, there is one dominant external IP, but the other endpoints are also present, which can be interpreted as the presence of either multi endpoints command and control (C2) or frequent web usage.

3. Distorted traffic distribution: Other internal hosts only add a few packets (in terms of packet counts), or bytes (in terms of byte counts), thus highlighting the abnormality posed by 147.32.84.165 to other hosts in comparison.

   The presence of traffic that can be concentrated due to one internal host is likely to suggest automated traffic, namely botnet traffic or heavy downloads. The behaviour of DNS, the patterns of the HTTP requests, and time of connection are to be examined more thoroughly to determine whether this phenomenon is harmless or reflects the malicious activities.

5. Traffic Analysis and Observations



ip.src == 147.32.84.165 && http

New Column	No.	Time	Source	Destination	Protocol	New Colum Lenç	Info
27	395	06:56:38.3234…	147.32.84.165	61.135.188.210	HTTP	645	GET /ie.png?02D%30%Dn3i%gsowlp6u2o%602D%31%Dr3a%BerpdydtAemgoarpFbiavtanNe6p2o%602D%30%Dt3n%enc
85	19198	06:59:52.3548…	147.32.84.165	123.126.51.33	HTTP	621	GET /images/entry_set_close.gif HTTP/1.1
87	19205	06:59:52.3593…	147.32.84.165	123.126.51.33	HTTP	618	GET /images/sync_loading.gif HTTP/1.1
83	19191	06:59:52.3542…	147.32.84.165	123.126.51.33	HTTP	618	GET /images/set_over_btn.gif HTTP/1.1
86	19203	06:59:52.3587…	147.32.84.165	123.126.51.33	HTTP	615	GET /images/save_btn0.gif HTTP/1.1
82	19183	06:59:52.3532…	147.32.84.165	123.126.51.33	HTTP	615	GET /images/save_btn0.gif HTTP/1.1
84	19196	06:59:52.3545…	147.32.84.165	123.126.51.33	HTTP	610	GET /images/add1.ico HTTP/1.1
78	1430	06:56:51.2951…	147.32.84.165	123.126.51.57	HTTP	570	GET /cl.gif?uigs_productid=iportal&rdk=1313495809837308&pars=sp%3A123%091%09pv%09123%09500%0945
75	1386	06:56:49.8326…	147.32.84.165	123.126.51.33	HTTP	534	GET /js3to1/sugg_ajaj_index.js?V=201108151717 HTTP/1.1
77	1390	06:56:49.8344…	147.32.84.165	123.126.51.33	HTTP	527	GET /js3to1/citydata.js?V=201108151717 HTTP/1.1
76	1388	06:56:49.8328…	147.32.84.165	123.126.51.33	HTTP	517	GET /images/entry_bg_add.gif HTTP/1.1
80	1535	06:56:55.7509…	147.32.84.165	61.135.188.210	HTTP	477	GET /pv.GIF?t=1313495815195989&u=1313495815195989&r=&pl=http://ie.sogou.com/features.html&load=
34	450	06:56:38.6382…	147.32.84.165	123.126.51.64	HTTP	456	GET /getentry.php?m=50F8290AC57E77F533C2A3F1B4AB39E4&from=mysql&city=unknown&method=ajaj&newzx=
35	452	06:56:38.6385…	147.32.84.165	123.126.51.65	HTTP	390	GET /indexbody_v23.js?v=201108151717&cbid=indexbody HTTP/1.1
21	330	06:56:37.6470…	147.32.84.165	123.126.51.33	HTTP	376	GET /css3to1/skin_.css?V=201108151717 HTTP/1.1
53	907	06:56:42.9479…	147.32.84.165	123.126.51.33	HTTP	375	GET /static/gouwu/images/newicon.gif HTTP/1.1
58	955	06:56:43.4562…	147.32.84.165	123.126.51.33	HTTP	374	GET /images/shopping/taobaologo.gif HTTP/1.1
71	1136	06:56:45.2950…	147.32.84.165	123.126.51.33	HTTP	373	GET /index_v23_2.js?V=201108151717 HTTP/1.1
52	866	06:56:42.6222…	147.32.84.165	123.126.51.33	HTTP	373	GET /images/shopping/searchbtn.gif HTTP/1.1

Figure 3

The overall traffic analysis of the botnet-capture-20110816-sogou.pcap explains the communication trends, protocols, and transactions between the host (147.32.84.165) and other external IP addresses. The main observations are as follows:

5.1 Host Communication patterns

The host 147.32.84.165 communicates with the several external IP addresses, 123.126.51.33, 61.135.188.210, 218.29.42.137, and 195.113. 232.73. Most of the traffic, both in packets and bytes is to 218.29.42.137:

- Packets: 18,886
- Bytes: 18 MB
- Bits/s (A → B): 29kbps
- Bits/s (B → A): 663kbps

The other external IP addresses which are significantly visited and include:
- 123.126.51.33 (1,317 packets, 786kB)
- 61.135.188.210 (120 packets, 13kB)

5.2 Traffic Characteristics

In the process of examining the packet capture, it was noted that all HTTP traffic was comprised of only GET requests that were directed to various external resources such as images, scripts and an assortment of contents. Examples of such requests will include things like /imagesindex14/d2.jpg, /sogouexplorerupgrade2.2.0.2070.exe, and /js3to1/suggajajindex.js. There were no DNS requests that made it to external IP addresses and this leads us to believe that domain name resolution had already occurred before the capture or that it had already connected with pre-resolved IP addresses. TCP flow analysis indicates an acute asymmetry, as the number of outbound packets is significantly higher than that of inbound ones (Packets A -

> B > Packets B -> A). This is typical of automated botnet traffic, which attempts to call out commands or enlist information.

5.3 Notable Traffic Behavior

There was repeated HTTP GET activity on multiple external servers, 123.126.51.33 and 61.135.188.210, which is symptomatic of sustained and persistent connections and is likely to be the host. Part of the traffic was the download of executable binaries (.exe), image files (.jpg, .gif) and JavaScript (.js), a trend that is congruent with malware-induced auto-transmission over HTTP. The traffic is of a sporadic nature with requests being sent every several milliseconds to seconds, which also proves the assumption of scripted or automated traffic over organic and human-driven browsing.

5.4 Summary

The evidence of botnet-type behavior exited host 147.32.84.165, which is a large amount of HTTP GETs to many foreign IP addresses. No corrupted packets were found and the most common protocol used was that of HTTP over TCP. Outbound packet excceds inbound packet count and a variety of external IP addresses are used to request content and hence support a traditional model of automated botnet communication.

6. Stream and top talker Analysis

From the capture, we can identify which external IPs received the highest volume of traffic from the host 147.32.84.165.

External IP	Total Packets	Total Bytes	Observations
218.29.42.137	18,886	17,675,752	Highest traffic volume; automated GET requests; likely command & control or data exfiltration
123.126.51.33	1,317	786,158	Multiple GET requests for images, scripts, and HTML pages; repeated access indicates automated browsing
147.32.84.165 → 61.135.188.210	120	12,787	GET requests for ie.png and other resources; smaller data payload but persistent communication
195.113.232.73	19	5,921	Occasional GET requests, could be update checks or secondary server communication
123.126.51.65	42	23,074	Similar pattern to 123.126.51.33; targeted HTTP GETs

6.2 Stream Analysis

The data lists the names of streams and directions of packets, which makes it possible to systematically analyze the conversation flows between the host and the third party.

Streams with high volume:

- ➔ Stream ID 14 (to 218.29.42.137) is shown to have received 12,557 packets between A and B along with 6,329 packets between B and A, which testifies to the ongoing automated traffic.
- ➔ Stream identifiers to 123.126.51.33 repeatedly send GET operations toward images, CSS files, and JavaScript assets, which is an indication of scripted web browsing or an automated download.

Packet direction insights:

- ➔ A toB (host to external) traffic is dominating, which means that most requests are initiated by the host.
- ➔ B to A (external to host) traffic is mostly made up of response traffic which includes, but is not limited to, HTML pages, images and scripts, and is relatively low in volume.

6.3 Patterns of Suspicious Behavior

There is repeated HTTP GET request to various external IPs which happens in quick succession. Such requests include executables (.exe), scripts (.js), and media files, which are the common signatures of a malware activity that pulls down payloads or command instructions. There is nothing in the logs indicating DNS queries, thus indicating that direct IP-based connections are used, an expected method to avoid typical network detection systems. Streams with a high volume of packets like the one between 218.29.42.137 can indicate data exfiltration or botnet command retrieval.

6.4 Summary

Host 147.32.84.165 becomes the main active host, trying to attack several external IP addresses using the GET requests. The traffic profile is dominated by principal talkers such as 218.29.42.137 and 123.126.51.33 which is a testament of automated behavior. The stream analysis supports the botnet-like communications, which are based on multistage outbound requests and a variety of payloads.

7. Conclusion

The analysis established the existence of network behavior that is compatible with botnet infection in the data under analysis. One of the internal systems exhibited atypical communication properties, such as persistent outgoing traffic, concentration of connections to particular external destinations, and the patterns of activity that were consistent with automated beaconing and not interactive user behavior. Although the inspection of direct payload was not conclusive to identify the malicious commands, the structural and statistical characteristics of the traffic were indicative enough to determine the host to be potentially compromised. The case illustrates the usefulness of behavioral network analysis in identifying command-and-control activity without signature-based detection and the significance of traffic baselining and anomaly detection in the enterprise threat hunting process.