TEAM CYMRU
COMMUNITY SERVICES

NEWS

SERVICES

· Bogon Reference
· Darknet Project
· IP to ASN Mapping
· Malware Hash Registry
  · WinMHR
· Open Resolver Challenge
· Screensaver
· BATTLE
· BIN Feed

MONITORING

READING ROOM

ABOUT US

Team Cymru Research NFP is a specialized Internet security research firm and 501(c)3 non-profit dedicated to making the Internet more secure. By researching the 'who' and 'why' of malicious Internet activity worldwide, Team Cymru helps organizations identify and eradicate problems in their networks.

Episode 64: Comics

MALWARE HASH REGISTRY

The Malware Hash Registry (MHR) project is a look-up service similar to the Team Cymru IP address to ASN mapping project. This project differs however, in that you can query our service for a computed MD5 or SHA-1 hash of a file and, if it is malware and we know about it, we return the last time we've seen it along with an approximate anti-virus detection percentage.

SPECIAL NOTICE

The Malware Hash Registry (MHR) is free for non-commercial use ONLY. If you wish to discuss commercial use of this service, please contact Team Cymru for more information.

If you are planning on implementing or automating the use of this service in any free or open software, application or host, PLEASE let us know in advance. We would like to adequately plan for capacity and make sure that we can handle the additional load you may generate. Please use the WHOIS-based service for larger queries. We have had instances where large deployments are put in place without informing us in advance, making it difficult to maintain a stable service for the rest of the community.

Attempting to enumerate the malware registry via the public service interface is not only impractical, it is also strictly prohibited. Contact us if the public interface is insufficient for your needs and we may be able to come up with alternative arrangement.

INTRODUCTION

Team Cymru is happy to announce the availability of various service options dedicated to mapping suspected malware hashes to our insight about positively identified malware. Now you can check if a particular piece of code is malware by querying against the extensive Team Cymru Malware Hash Registry.

The Team Cymru Malware Hash Registry (MHR) compliments an anti-virus (AV) strategy by helping to identify unknown or suspicious files. While your AV posture helps you perform detection based on signatures, heuristics and polymorphism, the MHR provides you additional layer of detection, for known badness. Based on our research, AV packages have trouble detecting every possible piece of malware when it first appears. The MHR leverages multiple AV packages and our own malware analysis sandbox to help aid your detection rate. Coupled with AV, the MHR helps identify known problems so you can take action. In order to decrease the false positive rate, we do not list items with less than 5% detection rate, we exclude all entries present in the NIST database, and we attempt to exclude multiple copies of polymorphic malware.

The service options come in various flavors, including:

- Whois (TCP 43)

- DNS (UDP 53)
- HTTP (TCP 80)
- HTTPS (TCP 443)
- WinMHR (Windows application - COMING SOON!)

Additional features are being considered for the future. Contact us with your ideas!

Following is a brief summary on how to use each of the services.

WHOIS

The whois daemon acts like a standard whois server would, but a MD5 or SHA-1 hash value instead of a name or address is passed as an argument. It accepts arguments on the command-line for single whois queries and it also supports BULK hash submissions when combined with GNU's netcat for those who wish to optimize their queries. When issuing requests for two or more hashes we strongly suggest you use netcat for BULK submissions since there is less overhead.

WARNING: Source addresses or networks that are seen abusing the whois server with large numbers of individual queries instead of using the bulk netcat interface may be null routed. Sources issuing an abnormally large number of queries may be automatically rate-limited. The netcat interface should be used for large groups of hash lists at a time in one single TCP query.

There is presently one whois server available with round robin IP addresses:

- hash.cymru.com

The syntax for whois and netcat whois IP queries is as follows:

```
Whois   Netcat          Action
        begin           enable bulk input mode          (
        end             exit the whois/netcat client    (
help    help            the help message
```

An example use of the command-line arguments on a single malware hash query:

```
$ whois -h hash.cymru.com e1112134b6dcc8bed54e0e34d8ac272
e1112134b6dcc8bed54e0e34d8ac272795e73d74 1221154281 53
```

The output above includes the hash that was queried for, along with the last known GMT timestamp associated with that hash in unix epoch, and the detection percentage across a mix of AV packages. If the malware hash is NOT in the database, the results will look something like this:

```
$ whois -h hash.cymru.com 1250ac278944a0737707cf40a0fbecd
1250ac278944a0737707cf40a0fbecd4b5a17c9d NO_DATA
```

We recommend the use GNU's version of netcat, not nc. (nc has been known to cause buffering problems with our server and will not always return the full output for larger malware hash lists). GNU netcat can be downloaded from http://netcat.sourceforge.net. This is the same as gnetcat in FreeBSD ports.

To issue bulk queries, follow these steps:

1. Create a file with a list of hashes, one per line. Add the word begin at the top of the file and the word end at the bottom. Example of list01:

```
begin
7697561ccbbdd1661c25c86762117613
cbed16069043a0bf3c92fff9a99cccdc
...
e6dc4f4d5061299bc5e76f5cd8d16610
end
```

2. Run the list through GNU netcat (NOT the venerable nc).

```
$ netcat hash.cymru.com 43 < list01 > list02
```

   The file list02 should now appear as:

```
# Bulk mode; hash.cymru.com [2009-11-12 19:39:50 +00
# SHA1|MD5 TIME(unix_t) DETECTION_PERCENT
7697561ccbbdd1661c25c86762117613 1258054790 NO_DATA
cbed16069043a0bf3c92fff9a99cccdc 1231802137 69
...
e6dc4f4d5061299bc5e76f5cd8d16610 1258054790 NO_DATA
```

Additional help can be obtained by issuing the help command:

```
$ whois -h hash.cymru.com help
```

DNS

The DNS daemon is designed for infrequent, but rapid lookups, much in the same way as other remote blackhole list (RBL) lookups are done. DNS has the added advantage of being able to cache answers locally and is based on UDP so there is much less overhead from a client perspective. Prepend the hash value as a label to the malware.hash zone:

- malware.hash.cymru.com

There are two types of queries you can perform, a TXT or an A query. The TXT query will give a bit more verbose output, including the last seen timestamp and an anti-virus package detection rate if the hash exists in our registry. If the hash exists in our registry and you just issue a default A query a loopback address will be returned. The address returned if a positive result is found should always be 127.0.0.2.

The format and output for a DNS TXT query is as follows:

```
$ dig +short 733a48a9cb49651d72fe824ca91e8d00.malware.hasl
"1221154281 53"
```

The format and output for a DNS A query is as follows:

```
$ dig +short 733a48a9cb49651d72fe824ca91e8d00.malware.has
127.0.0.2
```

If a given hash does not exist in our registry, the daemon will return a standard NXDOMAIN response (domain does not exist). If you have been rate limited, you will not receive any response and your packet will be dropped.

HTTP/HTTPS

The HTTP/HTTPS interface to WebMHR acts as a web-based proxy to the underlying WHOIS service. You can reach the web interface by browsing to:

http://hash.cymru.com/ or **https://hash.cymru.com/**

Simply follow the instructions on either of the pages at the links above to submit your hashes via the web.

FREQUENTLY ASKED QUESTIONS (FAQ)

1. How do I interpret the output?

   If a hash exists in our registry and is identified as malware there are two output values of interest. One is a timestamp when the malware was last seen, the other the rough anti-virus package detection rate.

   The timestamp is a UNIX time aka POSIX time whose value is the number of seconds since midnight January 1, 1970 universal coordinated time (UTC). So for example, 1223478925 seconds since midnight 1970-01-01 is Wednesday, October 8 15:15:25 UTC 2008. With a bash shell in unix you can map between UNIX time and to a more readable local time use the command `date --date="1970-01-01 <unix timestamp> secs UTC"`. Using Perl, you can use this command `perl -e 'print scalar localtime(<unix timestamp>),"\n"'`.

   The anti-virus package detection rate is a two or three digit value representing the total detection rate as a percent of all the anti-virus packages we ran against malware.

2. What anti-virus packages are you using?

   We try to use over 30 undisclosed anti-virus software packages. In a limited number of cases a smaller number of AV packages will be used.

3. What else can you tell me about your malware database?

   We collect and analyze every piece of malware we can get our hands on. More information will be available soon.

4. What anti-virus engine has the best detection rate?

   Some packages do a better or worse job on one particular

piece of malware, but it total we have not found any of them to reach the highest detection rates necessary to keep up with the ever increasing amounts of malware seen in the wild.

5. So, should I just not bother using an anti-virus package?

   NO! You need to realize their limitations, but you should definitely not abandon anti-virus packages or any other security solution, just because it is not 100% effective. Is anything? Use a defense in depth approach. Do your best to implement other security procedures and tools that will help in cases where malware makes it past your anti-virus defenses.

6. What anti-virus package should I use?

   It would be inappropriate for us to make a recommendation for a specific package, but we do suggest using a defense in depth strategy. Use multiple packages if you can. Use different types of packages, including signature-based and anomaly-based.

7. How do you collect malware?

   We employ various collection techniques such as honeypots, crawlers, and leverage private data sharing agreements with partners to aggregate it all to support our malware insight.

8. How up-to-date is your registry?

   The malware hash registry is reloaded once per day. Please note that we try to avoid including too much polymorphic malware when possible.

9. Can I have a copy of one or more piece(s) malware you have?

   Sorry, we believe it is inappropriate to share actual binary copies of malware with the general public. Additionally, many of our data-sharing agreement would not permit us to re-distribute samples.

10. Can I download your hash registry database?

    The hash registry database is not publicly available for download, but you may contact us about setting up a data sharing agreement should you need a more efficient method of performing queries.

11. Your service says my file is malware, but I know it is not!

    We're very sorry about that. We had hoped no one would have ever needed to ask this question, but if you have found a false positive we do want to know about it so we can fix a potential problem with our system. First, please be absolutely sure it is not malware. There are numerous free online services that will run your malware through multiple anti-virus packages. Virus Total is one such example. If you're sure the file is not malware, you will need to send us a copy of it. Use our main contact address

team-cymru@cymru.com and please encrypt the file with PGP to our public key in order to avoid any potential mail filtering problems.

12. This service is great! how can I help?

    We're glad you asked! Please contact us to let us know in what capacity you think you can be of assistance in terms of data or time. We're always looking for more partners.

13. I'm an AV vendor and I'd like to send you my anti-virus package, can you use it?

    Sure! Please send us the details regarding your AV package. Presently we only support Linux compatible CLI versions of AV products.

14. Can I send you malware?

    Please feel free to contact us as to how we may be able to receive your feed.

ADDITIONAL RESOURCES/TOOLS

We welcome contributions of tools that make use of the Malware Hash Registry, as long as it is used within the terms described above. If you make such a tool, please feel free to contact us with the details, and we may add it to the list here!

- The Malware Hash Registry and Bro-IDS (Seth Hall)

CONTACT US

Team Cymru | 16W361 S. Frontage Road | Suite 100 | Burr Ridge, IL 60527 | USA

Phone +1 630 230 5400
Fax +1 630 887 8651