MUDDLING MEERKAT:
THE GREAT FIREWALL MANIPULATOR

Authors:
Dr. Renée Burton
and Anonymous
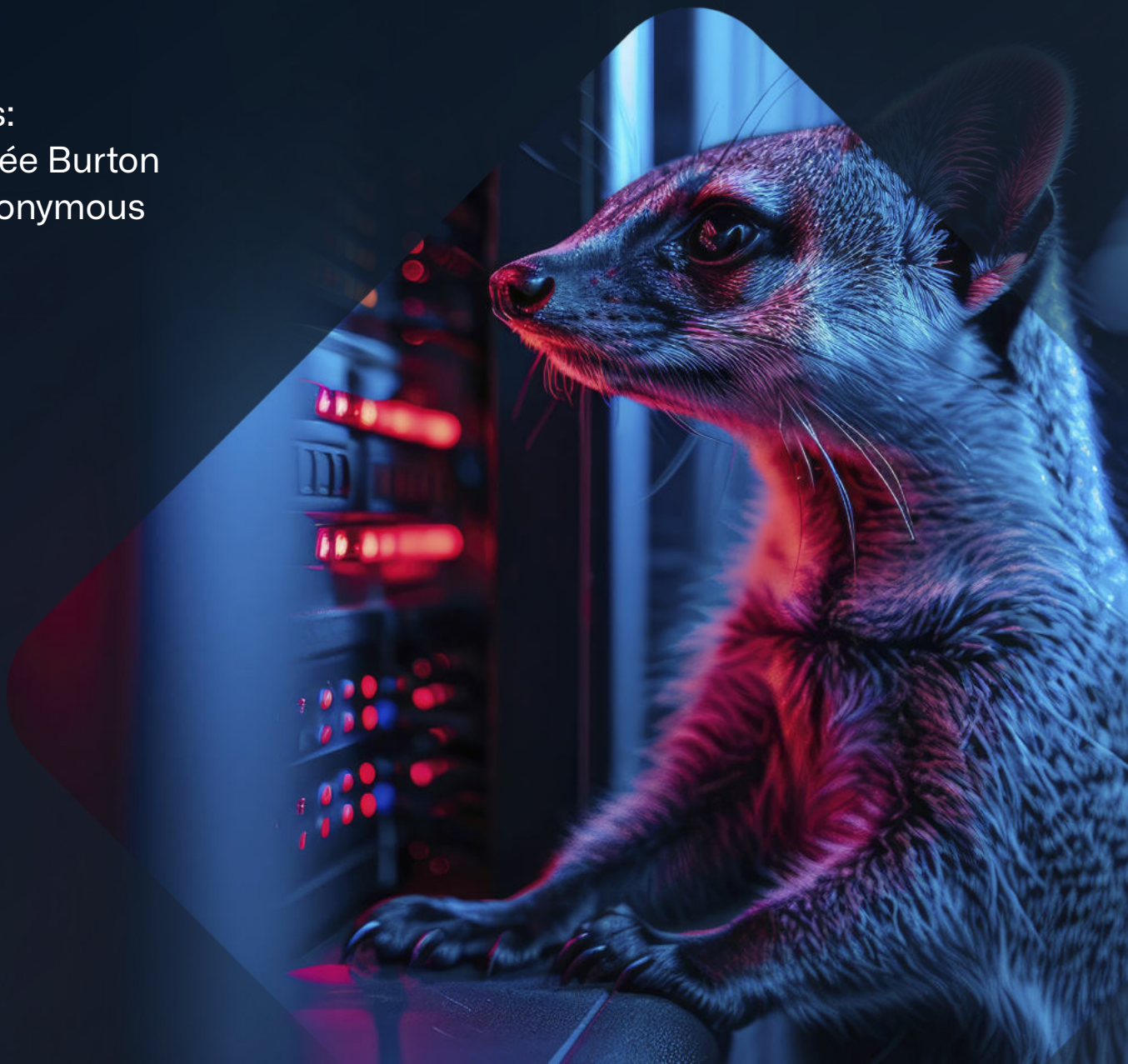
TABLE OF CONTENT

EXECUTIVE SUMMARY

This paper introduces a perplexing actor, Muddling Meerkat, who appears to be a People's Republic of China (PRC) nation state actor. Muddling Meerkat conducts active operations through DNS by creating large volumes of widely distributed queries that are subsequently propagated through the internet using open DNS resolvers. Their operations intertwine with two topics tightly connected with China and Chinese actors: the Chinese Great Firewall (GFW) and Slow Drip, or random prefix, distributed denial-of-service (DDoS) attacks. While Muddling Meerkat's operations look at first glance like DNS DDoS attacks, it seems unlikely that denial of service is their goal, at least in the near term. Muddling Meerkat operations are long-running — apparently starting in October 2019 – and demonstrate a high degree of expertise in DNS.

Muddling Meerkat's operations are complex. Indeed, they are so convoluted, one might assume that Muddling Meerkat presents no threat. But in cybersecurity, especially in the complex world of DNS, we should think strategically. In February 2024, the U.S. Cybersecurity and Infrastructure Security Agency (CISA) and several international partners issued an advisory saying, "In recent years, the U.S. has seen a strategic shift in PRC cyber threat activity from a focus on espionage to pre-positioning for possible disruptive cyber attacks against U.S. critical infrastructure."[1] While that specific advisory focused on "living off the land" techniques used by the actor Volt Typhoon, the message that "PRC cyber actors blend in with normal system and network activities, avoid identification by network defenses, and limit the amount of activity that is captured in common logging configurations" is eerily similar to how well-hidden Muddling Meerkat remains.[2]

WHAT IS MUDDLING MEERKAT

Muddling Meerkat has the apparent ability to control the GFW and does so in a way not previously reported. While parts of their operations are similar to Slow Drip attacks, the motivation and goal of Muddling Meerkat are unclear. The data shows us that their operations:

- Use servers in Chinese IP space to conduct campaigns by making DNS queries for random subdomains to IP addresses around the world, ultimately probing DNS networks globally

- Use MX record queries, plus other record types, for short random hostnames of a set of domains outside the actor's control in the .com and .org top-level domains (TLDs)

- Induce false MX records from Chinese IP addresses injected by the GFW

- Use "super-aged" domains, typically registered before the year 2000, avoiding DNS blocklists and colliding with many enterprise Active Directory domains

- Choose domains for abuse based on their length and age rather than their current status and ownership; while many of the domains are abandoned or have been repurposed for questionable use, other domains are actively used by legitimate entities

- Conduct campaigns of one to three days on a fairly continuous basis

- Do not appear to use large-scale spoofing of source IP addresses but instead initiate DNS queries from dedicated servers

- Are limited in size to avoid detection and service disruptions

- Are possibly conducted in discrete components, creating different DNS patterns over time

---

1   https://www.linkedin.com/posts/cisagov_with-us-and-international-government-partners-activi-ty-7161082451354603520-pv0q

2   https://www.cisa.gov/resources-tools/resources/identifying-and-mitigating-living-land-techniques

- Began on or about October 15, 2019[3]

A simplified view of Muddling Meerkat's operations as we understand them today is shown in Figure 1.
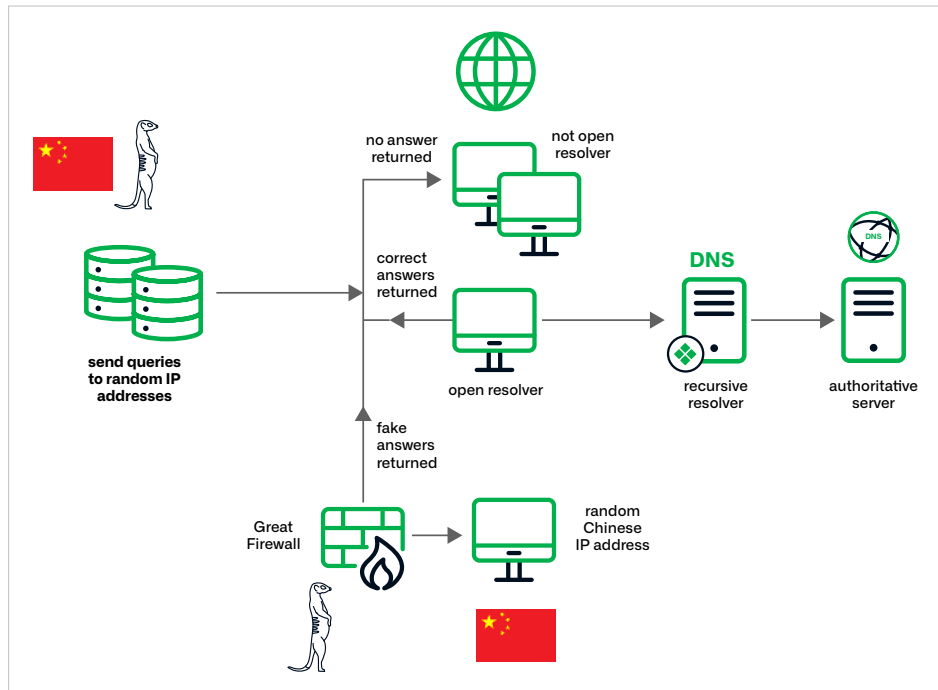

Figure 1. An overview of Muddling Meerkat operations as currently understood. The Great Firewall is observed providing fake answers to MX queries, a behavior that has not been previously documented.

Our discovery of Muddling Meerkat was serendipitous, and the actor could have gone undetected for many more years if not for the data visibility of multiple organizations. This paper is joint research with undisclosed threat researchers and security vendors, as well as the Merit Network, an independent non-profit corporation governed by Michigan's public universities, and DomainTools.[4] Each of the contributors has access to some form of passive DNS collection and can observe Muddling Meerkat from a unique perspective. It is impossible to observe the totality of Muddling Meerkat activities from any one vantage point. By combining information, we gain a picture of the actor's activity that would not be possible independently. Each finding within the paper, unless otherwise noted, is either confirmed in two independent sources or drawn directly from Infoblox DNS resolvers.

BACKGROUND

I've taken the unusual step of writing this paper in first person. In part, first person seems more appropriate when telling a strange tale like this. In addition, my prior studies and publications about Chinese DNS threat actors have helped inform my conclusions about Muddling Meerkat. Earlier in my career, colleagues at the National Security Agency (NSA) and I spent thousands of hours studying a Chinese actor who performed DNS-based DDoS attacks over several years. We dubbed that actor ExploderBot and quietly published those findings in the spring of 2018. After operating nearly daily since 2014, wreaking havoc on internet service providers, ExploderBot ceased operations just over a month after our paper was released. They have not been seen since May 18, 2018. The nature of the Chinese DNS

---

3   There is some evidence that the operations began a few months early, in June 2019, but I am unable to validate this date.

4   https://www.merit.edu/

DDoS attacks changed, and I wrote a longitudinal study on the changes in late 2020. Since then, I haven't spent much time looking at DNS DDoS attacks, Chinese or otherwise. We have detectors at Infoblox that look for signs of activity and automatically block the related domains for customers of our Advanced DNS Protection (ADP) product, but that system largely works with no need for human intervention.

Muddling Meerkat came to my attention while investigating a DNS threat actor that provides services for other threat actors dealing in illegal Chinese gambling and fake apps. It was not gambling that stood out but anomalous queries and responses for DNS mail server (MX) records. Though I found that Muddling Meerkat uses other record types as well, this paper will focus on MX records because their specific nature within DNS allows for cleaner analysis.

The GFW acts to prevent Chinese residents from accessing websites or services the government deems inappropriate or illegal.[5] But it is also known to inject false answers to DNS queries. The GFW applies to all IP traffic that crosses into, or out of, Chinese IP space. It is easy to demonstrate the GFW false answer behavior as I'll show later, in the section on Probing China's Great Firewall. The GFW can be described as an "operator on the side," meaning that it does not alter DNS responses directly but injects its own answers, entering into a race condition with any response from the original intended destination. When the GFW response is received by the requester first, it can poison their DNS cache. In addition to the GFW, China operates a system referred to as the Great Cannon (GC). The GC is an "operator in the middle," allowing it to modify packets en route to their destination.[6] The GC has been used for large-scale DDoS attacks. In 2015, it was used to attack the non-governmental organization GreatFire.org that monitors censorship at the GFW.[7] It has been used intermittently since then for DDoS attacks, including ones intended to prevent protests in Hong Kong.[8] The true scope of GC operations is unknown. In combination, the GFW and GC create a lot of noise and misleading data that can hinder investigations into anomalous behavior in DNS. I have personally gone hunting down numerous trails only to conclude: oh, it's just the GFW.

In addition to the abuse of MX records, Muddling Meerkat attracted our attention because it showed similar behavioral patterns, though at lower volumes, to DNS DDoS attacks. In a Slow Drip, or random prefix, DNS DDoS attack, queries for apparently random subdomains of a target domain are made on a large scale, typically propagated through open resolvers. These attacks originally emerged in 2014, and the first reported victims were Chinese. Several colleagues and I investigated DNS logs for multiple years of these attacks, concluding that most attacks that did demonstrable damage were conducted by a single actor, ExploderBot. We identified multiple mathematical artifacts in ExploderBot DNS queries and IP packets that remained consistent over five years. We also determined that the traffic from ExploderBot, which included spoofed source and destination IP addresses, was injected close to the internet backbone. Open resolvers that received the queries would forward them to their own recursive resolver, and in networks with many unmanaged devices containing unknown open resolvers, the query volume would disrupt internet server providers. The spoofed IP addresses used in ExploderBot DNS queries were broadly distributed, and the GFW responses served as red herrings hindering our analysis for a long time. When ExploderBot operations ceased in May 2018, what remained was a curious set of ongoing low-volume attacks with little apparent impact or purpose. In the past few years, random prefix attacks have impacted name servers somewhat regularly, but I have not seen the same volume level associated with ExploderBot.[9]

---

5   https://www.cybereason.com/blog/malcious-life-podcast-the-great-firewall-of-china-part-1

6   https://citizenlab.ca/2015/04/chinas-great-cannon/

7   https://foreignpolicy.com/2015/04/10/great-cannon-china-internet-cyber-attack-baidu/

8   https://cybersecurity.att.com/blogs/labs-research/the-great-cannon-has-been-deployed-again

9   https://infosec.exchange/@ricci@discuss.systems/111508151184559310

In this paper I will describe Muddling Meerkat operations in the context of what I know about the GFW, explain how to detect their activity and discuss some of the pitfalls of trying to analyze actors like Muddling Meerkat. In particular, I want to warn readers about the dangers of open resolvers and the use of unregistered search domains in DNS or Microsoft Active Directory, which can lead to both participation in DDoS attacks and leaking network information to adversaries.

A LITTLE LINGO

Language in DNS is confusing. When we compound it with IP packets, it becomes even more so. Several times in the course of this research, my coauthor and I had to stop and ask ourselves: *what IP are we talking about here??* Here is how I use several terms throughout the paper:

- The IP address that makes a DNS query, or receives a response for a DNS query, is called the **querier IP address**. This name applies whether the IP packet contained the query or the response.

- The IP address that responds to a DNS query is called the **responder IP address**. In a perfect world, these are resolvers, but as we'll see later in the section entitled The Role of Chinese IP Addresses, with Muddling Meerkat, they are just IP addresses.

- An IP address included in a DNS resource record of a response is called a **resolution IP address**.

- When I talk generally about DNS resource records in a response, I might say the **answer** refers to the value(s) contained in the record.

MUDDLING MEERKAT OPERATIONS

Muddling Meerkat operations are complex and demonstrate that the actor has a strong understanding of DNS, as well as internet savvy. To simplify this exposition, I cover only those components of the operation related to DNS MX records or MX resolution chains. In all cases, there is a registered domain, *not* under the control of the actor, called the **target domain**. I discuss three types of activity in this paper:

- Queries for MX records of a target domain
- Queries for MX records of random hostnames of a target domain
- Queries for A records of random hostnames of a target domain

Queries for random hostnames of a target domain typify a Slow Drip DDoS attack; however, Muddling Meerkat queries differ from those in ExploderBot or other Slow Drip attacks. The hostnames are short. Additionally, while some Slow Drip attacks do include a range of query types, the most common type is still an A record for an IPv4 address. I have not previously seen the type of MX record activity that characterizes Muddling Meerkat. The choice of target domains is also notable, as we'll see later in the Muddling Meerkat Target Domains section.

As for the name Muddling Meerkat: The meerkat is a member of the mongoose family. Deceptively cute in appearance, it is clever, industrious and exceptionally ferocious for its small size. Muddling Meerkat is known to abuse MX DNS records and conduct operations that involve the Chinese Great Firewall, adding confusion and red herrings to foil analysis. Due to the broad use of open resolvers for the operation, the activity also "pops up and down" over time and location, as meerkats do from their burrows.

PROBING CHINA'S GREAT FIREWALL

The GFW plays an important role in Muddling Meerkat data in that we can observe false responses to DNS queries in select DNS data collections. When we see a false response, the source IP of that record is a Chinese IP address, consistent with injection by the GFW or modification by the GC. Second only to the United States, China controls over 350 million IP addresses, geographically distributed around the world. For all traffic going into and out of this IP space, the GFW can inject answers to DNS queries using secretive decisions and without performance impacts to the user. To do this well requires a lot of expertise. China leveraged Western technology companies at the turn of the century to build components of the firewall and implement various other surveillance mechanisms, and in doing so, it built up its own capabilities and knowledge.[10]

China engineered a system that will respond with false answers rather than simply using an NXDOMAIN or other response mechanism that DNS firewalls commonly use.[11] Because of this, you don't need to take my word for it; you can probe the firewall yourself. Researchers have previously found false responses for hundreds of thousands of domains and concluded that some of these responses had polluted the cache of certain recursive resolvers.[12] In my research, both in that published on ExploderBot and since then, I've seen a dizzying array of IP address responses from the GFW.

The easiest way to demonstrate the impact of the GFW is to make DNS queries to a random Chinese IP address, one that is not an established DNS server. Stephen Bortmeyer provided a description of this in a 2015 blog.[13] Experiments can be done from the command line with the dig utility or with an online tool. If you ask for the A record of a popular domain, the Chinese IP address will invariably return an answer, even though it hosts no DNS service. Figure 2 below shows an example in which an IP address assigned to China Unicom and currently hosting no services, responds to a DNS query for the IP address of google[.]com with a fake answer.

```
; <<>> DiG diggui.com <<>> @111.193.204.201 google.com A
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 54398
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;google.com.                    IN      A

;; ANSWER SECTION:
google.com.             60      IN      A       93.46.8.90

;; Query time: 214 msec
;; SERVER: 111.193.204.201#53(111.193.204.201)
;; WHEN: Tue Jan 09 00:15:06 UTC 2024
;; MSG SIZE  rcvd: 54
```

Figure 2. A China Unicom IP address that hosts no services responds to DNS queries for the A record of google[.]com *with an IP address in Italy. The response is a purposeful redirection and will change in each response. Image credit:* diggui[.]com.

---

10 https://www.cybereason.com/blog/malcious-life-podcast-the-great-firewall-of-china-part-1

11 https://citizenlab.ca/2021/11/gfwatch-a-longitudinal-measurement-platform-built-to-monitor-chinas-dns-censorship-at-scale/

12 How Great is the Great Firewall? Measuring China's DNS Censorship. Nguyen Phong Hoang, et al., 30th USENIX Security Symposium (USENIX Security 21), https://www.usenix.org/system/files/sec21-hoang.pdf (last accessed Jan. 9, 2024)

13 https://www.bortzmeyer.org/sichuan-pepper.html

It is unknown how the GFW chooses which domains to send fake responses for as a means of censorship. Querying the same Chinese IP address for an uncensored domain will typically result in an error that no server could be reached. This result demonstrates that the GFW injects answers only for certain queries. In my experience, the GFW answers all DNS queries, regardless of the requested resource type with an IPv4 address. For example, if we ask the same IP address for the MX record of google[.]com, it returns a different IPv4 address, this time assigned to Korea Telecom. A proper MX record should include a text string with a fully qualified domain name (FQDN), not an IPv4 address. (See Figure 3.) A query for a TXT record or other non-A record type would similarly return an IPv4 address. Other researchers conducted large-scale longitudinal studies on the GFW in 2021 and reached the same conclusion.[14] A year earlier, a different set of researchers reported a single instance of a CNAME record injection, but they did not describe the response.[15]

```
; <<>> DiG diggui.com <<>> @111.193.204.201 google.com MX
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 62080
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;google.com.                    IN      MX

;; ANSWER SECTION:
google.com.            60       IN      A       59.24.3.174

;; Query time: 208 msec
;; SERVER: 111.193.204.201#53(111.193.204.201)
;; WHEN: Tue Jan 09 00:25:37 UTC 2024
;; MSG SIZE  rcvd: 54
```

Figure 3. A China Unicom IP address returns a random IPv4 address in response to an MX query for google[.]com. *A correct response would return the FQDN of the mail server. Image credit:* diggui[.]com.

These experiments show firsthand how the GFW typically operates. It selectively injects DNS responses for certain domain names with random misleading answers. When it inserts fake packets, it always returns an IPv4 address regardless of the requested record type. Muddling Meerkat, on the other hand, serves properly formatted fake MX records from Chinese IP addresses.

MX RECORDS FOR A TARGET DOMAIN

The most remarkable feature of Muddling Meerkat is the presence of false MX record responses from Chinese IP addresses. This behavior, never published before, differs from the standard behavior of the GFW. These resolutions are sourced from Chinese IP addresses that do not host DNS services and contain false answers, consistent with the GFW. However, unlike the known behavior of the GFW, Muddling Meerkat MX responses include not IPv4 addresses but properly formatted MX resource records instead. This feature is truly remarkable and largely inexplicable.

I'll use one of the many Muddling Meerkat target domains, kb[.]com., to demonstrate their activity throughout this paper. The MX answer records for Muddling Meerkat are only observable in data collected outside of the normal DNS resolution chain because the source of the response is not a DNS resolver but instead a random Chinese IP address. Because Infoblox data is derived from our recursive resolvers, I partnered with other vendors to obtain data for analysis.

---

14   How Great is the Great Firewall? Measuring China's DNS Censorship. Nguyen Phong Hoang, et al., 30th USENIX Security Symposium (USENIX Security 21), https://www.usenix.org/system/files/sec21-hoang.pdf (last accessed Jan. 9, 2024)

15   Anonymous, et al. Triplet Censors: Demystifying Great {Firewall{\textquoteright}s} {DNS} Censorship Behavior, 10th USENIX Workshop on Free and Open Communications on the Internet (FOCI 20), https://www.usenix.org/conference/foci20/presentation/anonymous (last accessed Jan. 9, 2024)

One third party provided DNS query-response data containing MX resource records for the domain `kb[.]com` over a period of 120 days ending in late January 2024. Specifically, each log included a DNS query for the MX record of `kb[.]com` and a response containing two resource records. The resource records were properly formatted, containing FQDNs with random hostnames of `kb[.]com`, typically three to six characters long. Examples of such MX record values include:

- `pq5bo[.]kb[.]com`

- `uff0h[.]kb[.]com`

- `biuti[.]kb[.]com`

- `8jxg1x[.]kb[.]com`

- `8p0[.]kb[.]com`

For those not familiar with MX records, these responses should be the FQDN of the mail server for `kb[.]com`. In order to deliver mail from a user on a network to a recipient in the `kb[.]com` network, two DNS queries are necessary. The first is for the MX records of the receiver's mail domain, here kb[.]com, and the second is for the IP address of the FQDN contained within the MX record. Once the IP address is obtained, the Simple Mail Transport Protocol (SMTP) server can send mail on the behalf of a user. (See Figure 4.)
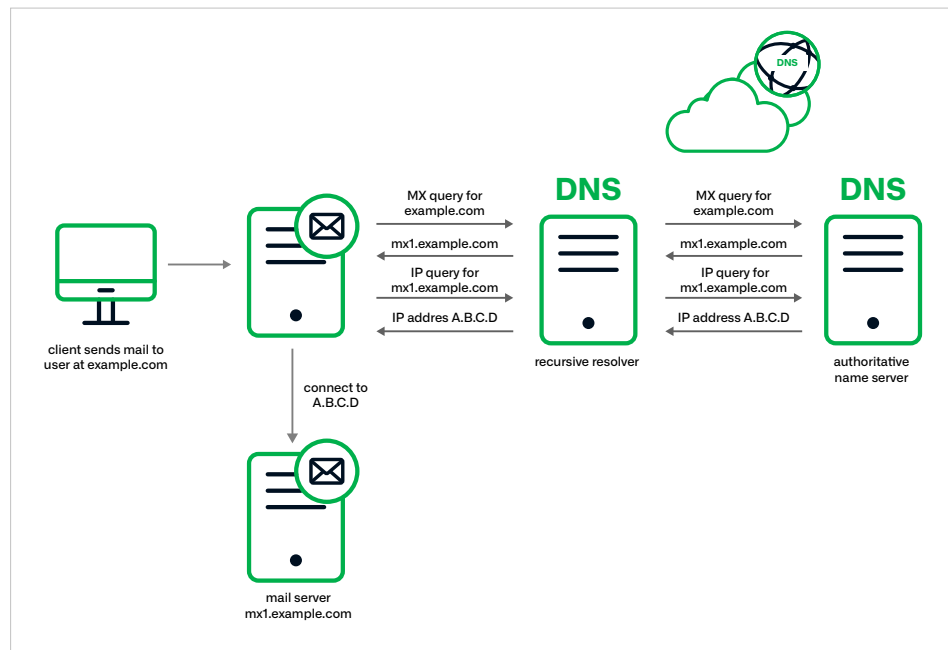

Figure 4 . The typical DNS resolution process to find a mail server IP address In the standard resolution for a mail server queries for both an MX record and an A record will occur.

In the third-party data, properly formatted MX records are sourced from random Chinese IP addresses that do *not* host DNS servers. Moreover, these answers, while appearing correct at first glance, are false. The domain `kb[.]com` currently has authoritative name servers in China with NS1, an authoritative name service that is part of IBM. These authoritative name servers return no response to MX record queries for `kb[.]com`. Thus, we observed DNS responses coming from Chinese IP space that both differed from the normal GFW behavior and were false.

The third-party data contained not just a few MX records, but thousands. Every hostname within the historical MX record set was seen on a single day during this time frame for a total of over 8k unique FQDNs. A second vendor has similar observations. The answers contain short hostnames and are not duplicated. The volume is notable but fairly small, certainly too small to be effective in DDoS attacks. Not only are the answers false here but the queries themselves also are suspect. The domain kb[.]com was once held by a U.S. marketing firm, but it now hosts geo-fenced Chinese language gambling. There is no reason for clients to send mail to the domain, and especially no reason to request resolutions from random Chinese IP addresses. As Figure 5 shows, there are MX resolutions for every day in the sample, but there are rarely more than 100 observations per day.
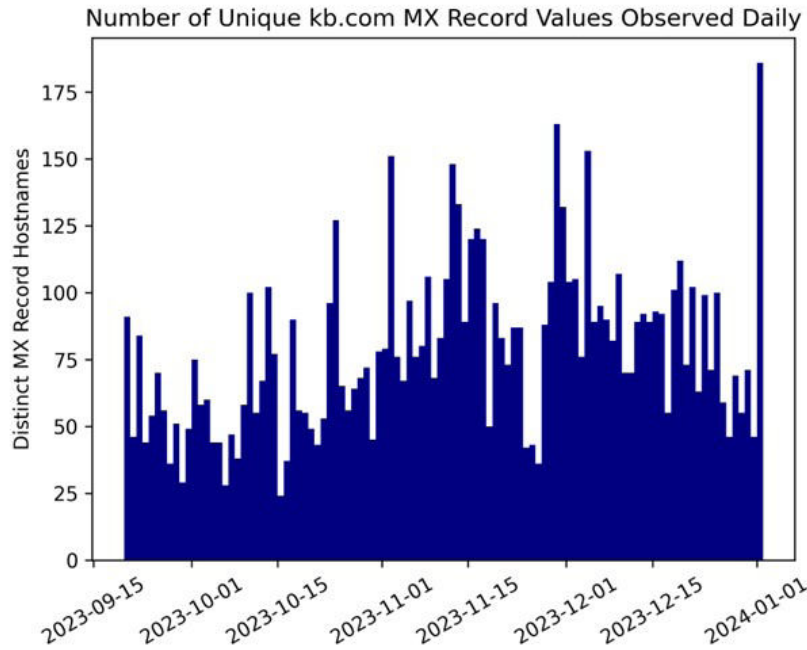

Figure 5. The daily count of unique MX record values for kb[.]com in the global pDNS collection. These are fake MX records that do not exist in the domain zone file.

We also analyzed historical answers for MX records of kb[.]com over several years (Figure 6). MX records containing a random hostname were first observed on October 15, 2019. We have independently verified with other vendors that the first MX resolutions for Muddling Meerkat target domains were first seen on, or about, October 15, 2019. This is true for all of the target domains we analyzed. Overall in third-party data, we see an inexplicable rise in the number of MX resolutions starting September 20, 2023, and continuing into early 2024.
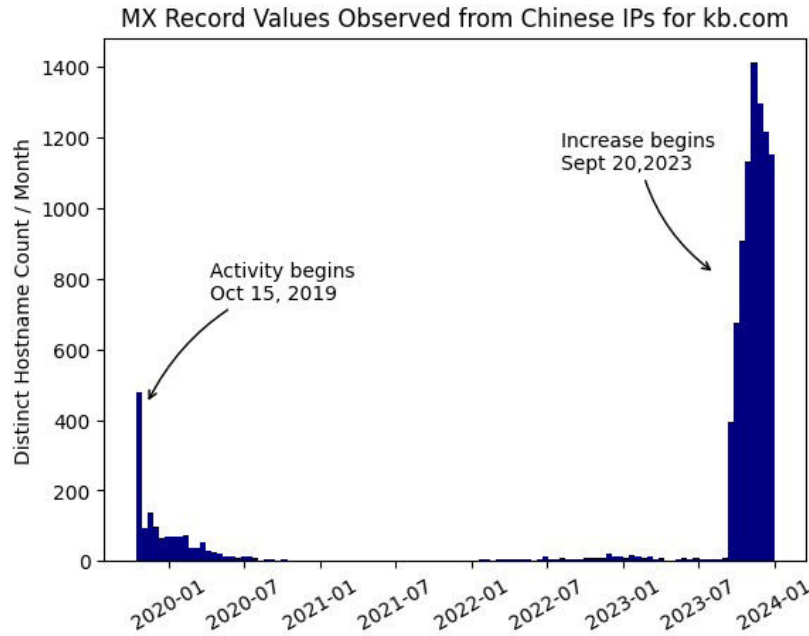
Figure 6. Count of unique fake MX record values for kb[.]com, *aggregated monthly over time and observed in third-party DNS data collections. The answerer IP addresses for these resolutions are random Chinese IP addresses which do not host DNS services, implying that the answer comes from the Great Firewall. These are all fake MX records that do not exist in the* kb[.]com *DNS zone file.*

A recursive resolver or other server along a normal DNS resolution path is unlikely to have seen these responses. Comparing the entire history of MX records for kb[.]com from both Infoblox and DomainTools Farsight, we have seen only a handful of unique records. As of January 2024, the name server for kb[.]com does not answer MX record requests from our resolvers. In the past, the authoritative servers have returned answers containing these values:

- mail.kb[.]com, smtp1[.]com, smtp2[.]com, smtp3[.]com

While the authoritative name servers for kb[.]com do not answer MX queries through the official DNS resolution process, our recursive resolvers do receive requests for these records. Under normal circumstances, receiving these requests would imply that users within our customer networks need to send email to a user at kb[.]com. But kb[.]com doesn't serve mail. Passive DNS logs contain many strange things, and queries can be triggered by old applications or websites. However, in this case, the queries occur exactly one month apart over several months, lending to the intrigue. As we will see from other data in the next section, this behavior is most likely triggered by Muddling Meerkat probing our customer networks for open resolvers and occasionally finding some.

I have been unable to manually trigger fake MX responses from the GFW, for Muddling Meerkat target domains or others. Perhaps the records are produced instead by the GC or in a specific Muddling Meerkat operational context. For example, the responses might be triggered by signatures within the IP packet that identify the actor. We know that ExploderBot IP packets contained multiple artifacts that could serve as a check on the source, if desired. The appearance of such identifying traces might explain why other researchers saw CNAME injections but only rarely. Unfortunately, this is all speculation based on prior experience and possible explanations for aberrant behavior by the GFW/GC. While the responses themselves could be fake IP packets, Occam's razor points to a variant of the GFW, possibly the GC. Many things are possible, but few are plausible.

MX RECORDS FOR A RANDOM SUBDOMAIN

The second identifying component of Muddling Meerkat operations also involves MX record queries—but for a random subdomain of the target domain, rather than the base domain itself. In this event, under normal circumstances, the query would be triggered by a user wanting to send email not to the base domain but to a subdomain. While this scenario does happen in normal DNS, it is not particularly common. In most of the Muddling Meerkat target domains, there is no functional mail server, creating an even more anomalous situation. Indeed, queries for MX records of random subdomains of kb[.]com are what led to this entire investigation.

The phenomena we observe at our recursive resolvers are a small number of requests occurring over one to three days with random hostnames. These requests include other query types besides MX records, but because of the specific nature of MX records in normal network operations, I am only reporting findings on this type. The MX queries have this form:

`<random>.target_domain`

where `random` is an alphanumeric string of variable length, typically between three and six characters long.

While this investigation began with kb[.]com, there are about 10 Muddling Meerkat target domains observed in our customer networks since September 1, 2023. Figures 7 and 8 show the volume of MX queries for kb[.]com and 4u[.]com seen at our recursive resolvers between September 1 and December 31, along with some sample FQDNs queried on specific days. Over this four-month period, no subdomain is repeated. Our partners at DomainTools Farsight and other undisclosed vendors observe the same trends, albeit with different random subdomains.
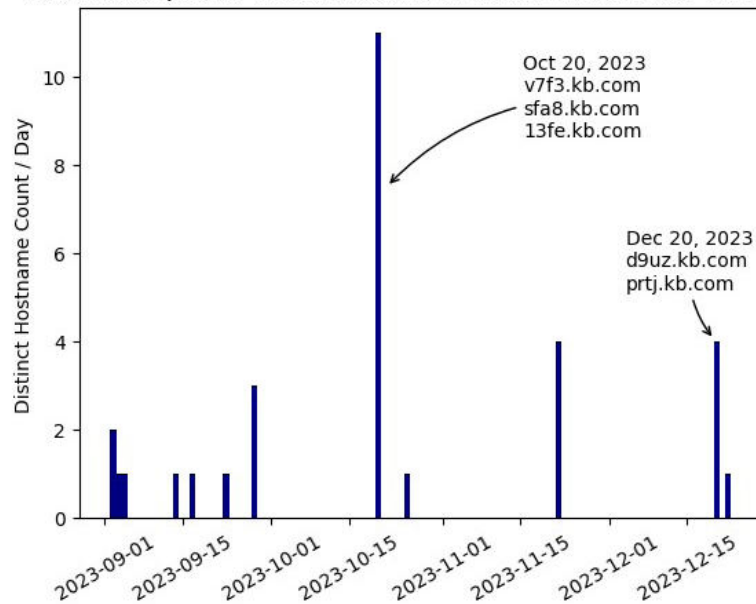

Figure 7. The number of distinct FQDNs with MX record queries for kb[.]com *seen at Infoblox recursive resolvers during four months*
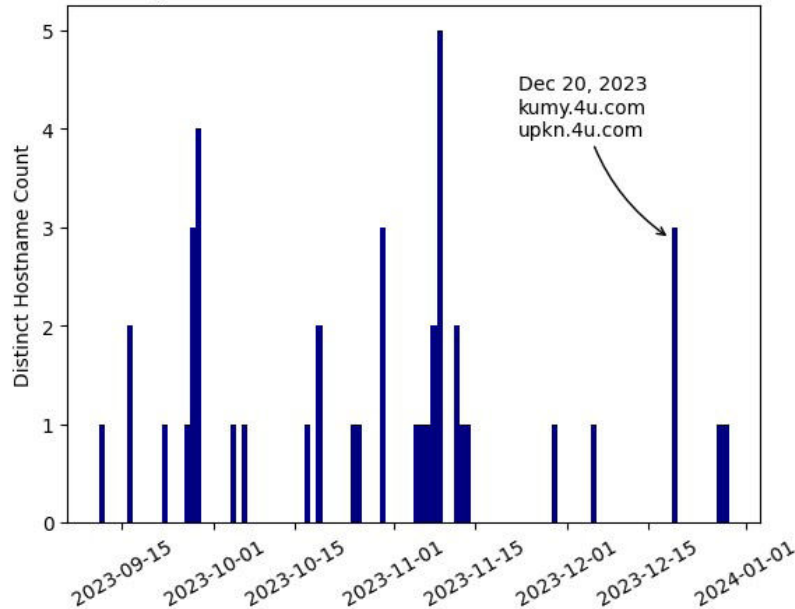
Figure 8. The number of distinct FQDNs with MX record queries for `4u[.]com` *seen at Infoblox recursive resolvers over a four-month period*

Figures 7 and 8 demonstrate the aperiodic "pop up" nature of Muddling Meerkat queries with an operational tempo that lasts one to three days and uses random hostnames. This kind of pattern is typical of Slow Drip DDoS attacks in general and ExploderBot specifically. However, there are some significant differences between what was previously reported in the literature and these attacks. Most notably, in these attacks, the volumes are much lower than we would expect for a real attempt at a DDoS and those seen in large-scale attacks at the height of this activity between 2014 and 2017.

In a longitudinal study published in the journal *Digital Threats Research* and Practice in 2019, I noted that the Slow Drip DDoS landscape had changed significantly since our first paper on ExploderBot.[16] In that research, conducted over six months in 2018, several query types were observed, but MX was not one of them. The dominant patterns described in that paper are still observed today, with low levels of queries with long hostnames and strong bias in character distributions. Muddling Meerkat has no similarity to those trends.

IPV4 RECORDS FOR RANDOM SUBDOMAINS

In addition to MX queries for random subdomains of the target domain, our recursive resolvers receive requests for A records, or IPv4 addresses. Of course, these queries do not receive answers from our resolvers because there is no such subdomain configured at the authoritative name server. Other vendors whose collection comes from recursive resolvers have similar observations. DomainTools Farsight data, for example, comes from a collection of recursive resolvers globally. Like Infoblox, those vendors see regular spikes in queries for random subdomains of the Muddling Meerkat domains, including A record queries. Figure 9 shows these trends for one month, January 2024.

---

16 Renée Burton. 2018. Unsupervised Learning Techniques for Malware Characterization: Understanding Certain DNS-based DDoS Attacks. Digit. Threat. Res. Pract. 37, 4, Article 111 (August 2018), 27 pages. https://dl.acm.org/doi/10.1145/3377869
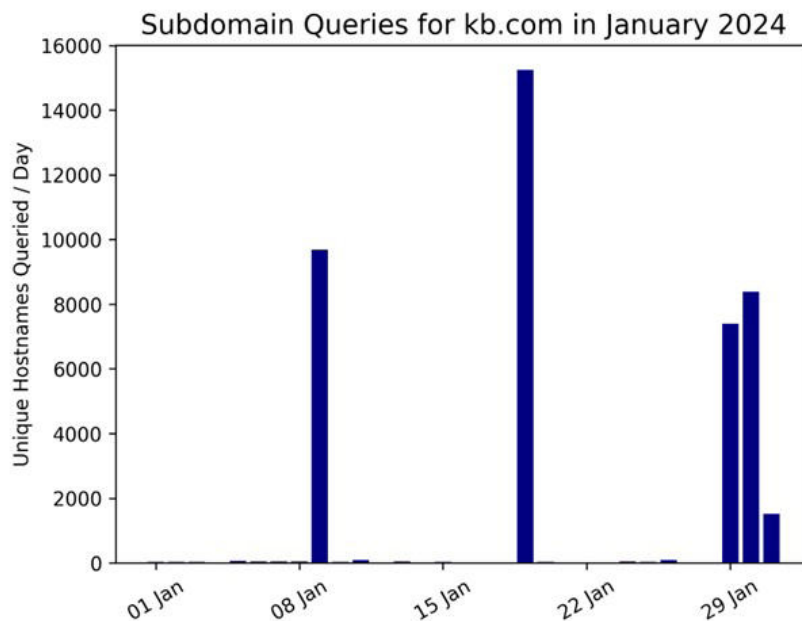
Figure 9. Unique hostname queries of kb[.]com *observed in Farsight pDNS in January 2024*

There are also other types of collection with visibility into DNS, including packet collection, honeypots, and internet telescopes. Working on the theory that the source of these queries within our networks was open resolvers, and that Muddling Meerkat likely was probing a broad spectrum of IPv4 space for open resolvers, I asked other vendors to help locate packets that contained resource records in the response. We found A record responses, just as we found MX record responses.

The only IP addresses that answered queries for A records of Muddling Meerkat domains were in Chinese IP space. These IP addresses were not open on port 53, meaning they were not DNS resolvers. In other words, these answers came from the GFW and not the authoritative servers.

The GFW is known to inject answers to DNS queries with resolution IP addresses that are not entirely random. In a longitudinal study covering nine months and published in August 2021 for the 30th Usenix Security Symposium, researchers found that the forged IP addresses often appeared repeatedly for certain groups of domains.[17]

Using IP resolutions of subdomains of kb[.]com, we mapped the occurrence of a forged resolution IP address with the timeline of queries. In every case, the resolution IP address is seen repeatedly, with distinct time windows lasting one to three days, for short random subdomains. Figures 10 and 11 show two examples of this behavior. The two IP addresses are not actually related to kb[.]com; these are fake answers from the GFW. Both IP addresses are seen on overlapping days. Each figure shows the entirety of resolutions for kb[.]com subdomains to that IP address in 2022. As with the Infoblox and Farsight resolver data, the hostname, or subdomain, is not repeated.

---

17 How Great is the Great Firewall? Measuring China's DNS Censorship. Nguyen Phong Hoang, et al., 30th USENIX Security Symposium (USENIX Security 21), https://www.usenix.org/system/files/sec21-hoang.pdf (last accessed Jan. 9, 2024)
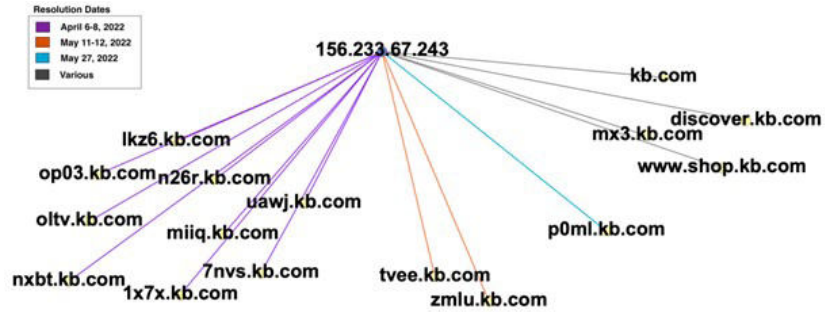
Figure 10. Hostname resolutions by the GFW within the kb[.]com *domain to the IP address 156[.]233[.]67[.]243 during 2022. This IP address is not related to* kb[.]com *and the answer is forged by the GFW.*
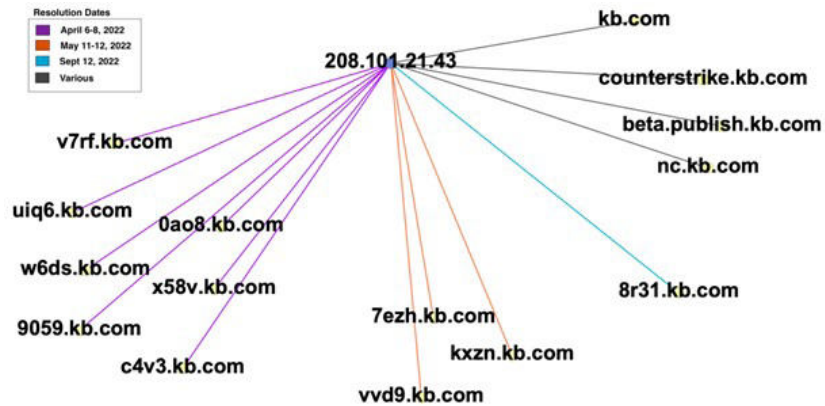

Figure 11. Hostname resolutions by the GFW within the kb[.]com *domain to the IP address 208[.]101[.]21[.]43 during 2022. This IP address is not related to* kb[.]com *and the answer is forged by the GFW.*

These results indicate that Muddling Meerkat is conducting operations that include DNS queries to a large number of destination IP addresses, regardless of their location or open ports, and that the GFW is injecting responses to these domains on specific days with a set of IP addresses that are used over time. This same activity and type of responses are ongoing in January 2024. While these figures show resolutions for kb[.]com, we have verified the same pattern for all of the known Muddling Meerkat target domains.

Here is where things get interesting: The GFW doesn't normally inject answers for kb[.]com or any subdomains. The GFW is not injecting fake responses to any random subdomain request of kb[.]com, only those created by Muddling Meerkat! As we discussed earlier, the GFW injects answers to popular domains or to domains that it finds somehow objectionable to Chinese interests. The aforementioned Usenix paper validates this fact. Figure 12 shows the response on January 13, 2024, to an A record query for nxbt.kb[.]com from the IP address 111[.]193[.]204[.]201 that we used earlier to get fake responses to google[.]com.

```
; <<>> DiG diggui.com <<>> @111.193.204.201 nxbt.kb.com A
; (1 server found)
;; global options: +cmd
;; connection timed out; no servers could be reached
```

Figure 12. The response to an A record request from 111[.]193[.]204[.]204 for nxbt[.]kb[.]com. *This IP address is in Chinese IP address space and is not open on port 53. The answer is what is expected for a query of this type and is consistent with known behavior of the GFW. Image credit:* diggui.com.

MUDDLING MEERKAT TARGET DOMAINS

The choice of Muddling Meerkat target domains demonstrates sophistication in DNS. Muddling Meerkat operators induce selective responses from the GFW that do not occur in normal GFW censorship. To do so, they have chosen target domains they do not control, which security appliances are very unlikely to block. Moreover, they use query types that are not commonly monitored and create a volume of queries that blends with normal DNS traffic. We have observed random hostnames with query types A (IPv4), CNAME, MX, and AAAA (IPv6) at Infoblox resolvers.

The random subdomain queries we have observed are for domains that have been registered for 20 years or more, have short labels, and are in the .com and .org TLDs. The target domain labels are mostly two or three characters long, but I have seen some examples that were four characters (e.g., boxi[.]com). In most cases, the domains have changed hands over time, but the original creation date will still be shown in WHOIS. Examples include kb[.]com, 4u[.]com, id[.]com, od[.]com, ntl[.]com, and nef[.]com. These domains were all observed in Muddling Meerkat traffic at Infoblox resolvers during December 2023 and January 2024.

I have verified approximately 20 target domains in multiple sources; however, there are likely many more. It is challenging to isolate the target domains for several reasons I will introduce here and discuss more in-depth later in the section entitled The Role of Chinese IP Addresses. First, not all domains that meet the basic age and length criteria appear to be targeted. For example, I have not found evidence that rr[.]com, ibm[.]com, and aol[.]com are used in Muddling Meerkat operations, although they meet the basic requirements. (*Yes, aol[.]com still occurs in DNS traffic.*) Most of the domains that are found in queries at our recursive resolvers are either not in use (e.g., 4u[.]com) or not particularly popular across customers. Many, like kb[.]com and od[.]com, are used for offshore Chinese-language gambling sites. A few, like ni[.]com, owned by National Instruments, are well-established, heavily used domains.

The choice to use long-established, short domains in well-reputed TLDs is clever for more reasons than the reduced probability of being blocked by security appliances. Domains with these characteristics are also frequently used:

• by organizations as DNS search domains or Active Directory domains and

• in malware to create red herrings for investigators

As a result, a security operations center (SOC) analyst who notices suspicious queries to these target domains will be stymied by the many potential sources of malware that might be connected to the query. For example, the domain kb[.]com has over 30 files referring to it and 7 files communicating with it, in samples stored by the vendor VirusTotal.[18] The domain od[.]com shows over 130 referring files.[19] Many of these are old malware samples and they add to the noise.

On the other hand, a researcher like myself, attempting to understand a more holistic picture of the activity, will have to filter through unrelated DNS queries to isolate the true target domains. This type of domain is commonly used for Active Directory by an organization, even though it does not control the domain. (*A risky practice!*) In addition, applications, websites, and humans cause aberrant queries in DNS. Of the range of query types used by Muddling Meerkat, MX is the easiest to analyze.

---

18   https://www.virustotal.com/gui/domain/kb.com/relations

19   https://www.virustotal.com/gui/domain/od.com/relations

To provide some perspective, I looked at MX resolutions at Infoblox recursive resolvers that occurred during six weeks starting December 1, 2023. When we think about mail server domains, we don't expect to see a lot of variety. But this expectation proves to be cognitive bias. I counted the number of SLDs with the following conditions that are similar to Muddling Meerkat:

• in the .com and .org TLDs

• result in NXDOMAIN responses

• have more than 10 different hostnames

More than 1,100 domains met the criteria. In short, lots of domains have anomalous MX queries. From those 1,100, I reduced the set to include only those where the domain label was less than four characters. This resulted in 55 candidates and over 22k unique queries during the study period. From this set of candidates, I conducted additional analysis to confirm target domains using a variety of other features.

THE ROLE OF OPEN RESOLVERS

An open resolver is a device on an IP address that will answer queries from any client, but it is not configured intentionally as a recursive resolver to serve the general public. In contrast, a public resolver in DNS is a recursive resolver designed to answer queries from any client and is typically run by a large business, such as Google, Cloudflare, or Yandex. Some researchers include public resolvers in their definition of open resolvers, but I do not. Open resolvers are well-known exploitation points for DDoS attacks. They can be used to amplify attacks against victims in reflection attacks, wherein DNS queries are made to open resolvers with spoofed sources containing the victim's IP address.[20] They are also used in Slow Drip attacks to distribute queries to the authoritative name server owned by the victim, and in variations of attacks against intermediate infrastructure.[21]

I use the term *IP address* here to describe open resolvers rather than a DNS resolver because open resolvers are very complex. For example, there may be an internet appliance, such as a firewall, in front of the open resolver IP address that can intercept queries and then, just like the GFW, forge a response, making it appear that the original destination IP address answered the DNS query. The answer returned may or may not be correct. This behavior is similar to that described by researchers on the interception of DNS queries by internet service providers (ISPs).[22]

Open resolvers both contribute to DDoS attacks and hinder analysis of them. They will create additional traffic to the root and TLD servers because they don't have the breadth of a DNS cache that a public resolver would have, frequently forcing them to perform a full resolution. In my experience analyzing open resolver traffic, many have other misconfigurations in their DNS, creating additional, typically unnecessary traffic. For example, they may not cache root hints and continually query for the root server IP addresses. When combined with the potential for forged responses, open resolvers create a lot of noise and produce red herrings for researchers.

20 A Matter of Degree: Characterizing the Amplification Power of Open DNS Resolvers, Yazdani, et al. Nature Switzerland AG 2022 O. Hohlfeld et al. (Eds.): PAM 2022, LNCS 13210, pp. 293–318, 2022. https://doi.org/10.1007/978-3-030-98785-5_13. https://annasperotto.org/publication/papers/2022/yazdani-pam-2022.pdf (last accessed Jan. 14, 2024)

21 NRDelegation Attack: Complexity DDoS Attack on DNS Recursive Resolvers, Yehuda Afek, et al., 32nd Usenix Security Symposium, 2023 https://www.usenix.org/conference/usenixsecurity23/presentation/afek (last accessed Jan. 14, 2024)

22 Who is Answering My Queries: Understanding and Characterizing Interception of the DNS Resolution Path, Baujun Lui, et al., 27th Usenix Conference, 2018 https://www.usenix.org/conference/usenixsecurity18/presentation/liu-baojun (last accessed Jan. 14, 2024)

I first encountered open resolvers while studying ExploderBot DDoS attacks. In these attacks, IP packets containing DNS queries for random subdomains of a target domain were dropped onto the internet near the backbone at various locations. Both the source and destination IP addresses were forged and, when taken together over time, covered a large amount of the IPv4 address space. In our research, we encountered all the aforementioned problems, including forged responses from the GFW and open resolvers. ExploderBot conducted these attacks typically over a few days, but they were aperiodic. Prior to 2016, there were operations many times a month, but these slowed in subsequent years and became very irregular. While ostensibly a DDoS attack on an authoritative name server, the greatest damage that ExploderBot caused was to ISP infrastructure, including recursive resolvers and load balancers. Without open resolvers, ExploderBot attacks would not have been notable, but for several years, although not attached to an actor name, their activity was covered in blogs and media reporting. ExploderBot is believed to be inactive; activity was last seen at Infoblox on May 18, 2018.

Open resolvers also play an important role in Muddling Meerkat operations. Evidence suggests that the queries are sent to a wide range of IP addresses, many of them open resolvers, from Chinese IP space. The destination IP addresses for the DNS queries likely rotate over time, which creates a "pop up" signature at recursive resolvers like Infoblox. In other words, I suspect Muddling Meerkat is actively muddling with the internet more often than we observe at Infoblox cloud resolvers. Instead, I suspect at certain intervals, lasting a few days at a time, external IP addresses belonging to our customers are included in the Muddling Meerkat destinations. (This is speculation on my part; I don't have data visibility to see the full scope of activities.) Some of our customers unwittingly have open resolvers in their network that receive their queries and forward them to our resolvers for resolution. Regardless of the operational tempo, we will only see Muddling Meerkat queries at our resolvers when a customer device forwards them.

Muddling Meerkat abuses many open resolvers. Some are established servers in a data center while others are home routers. For example, we observed a number of IP addresses that are fingerprinted as MikroTik routers by Shodan.[23] In January 2024, these IP addresses included queries from the sample open resolvers in Table 1.

Querier IP Address	Query Name
23[.]173[.]112[.]115	92ac[.]kb[.]com, mi2w[.]kb[.]com, 3k04[.]kb[.]com
103[.]47[.]134[.]195	zve3[.]kb[.]com, rjlf[.]kb[.]com, mayf[.]kb[.]com
38[.]54[.]105[.]163	q0ce[.]kb[.]com, h5ow[.]kb[.]com, 4e5r[.]kb[.]com

Table 1. Sample querier IP addresses and queries observed in January 2024; these IP addresses were all hosting open resolvers as of January 31, 2024

NO SPOOFED QUERIERS

Because of my experience with ExploderBot, I was predisposed to think that Muddling Meerkat was injecting DNS queries onto the internet using spoofed querier IP addresses and a broad spectrum of recipient IP addresses. The evidence we uncovered, though, indicated otherwise: Select Chinese IP addresses were the source of a disproportionate number of DNS queries. Based on the data (see Table 2 for examples), it seemed more likely that Muddling Meerkat was using dedicated servers for their operations.

23 Shodan.io is a publicly available search engine for server attributes by IP address.

In spite of the counter evidence, we wanted to test the spoofed querier hypothesis. The best way to do so is through what is called a **network telescope**,[24] which takes advantage of unused IP addresses to which there should be no traffic, and collects packets that are routed to them. Network telescopes are useful for capturing large-scale events that leverage spoofed IP addresses. A number of telescope operators, including Merit Network, are able to observe traffic to approximately 11 million IP addresses. Even though these IP addresses are technically unused, they receive a tremendous amount of traffic containing a wide variety of protocols.

In the context of a spoofed DNS query, the chain of events would go something like this:

- The attacker injects an IP packet that contains a DNS query purportedly from IP address A and directed for IP address B.

- Assuming IP address B is a DNS resolver, or an invisible proxy like the GFW, a response packet is sent from B to A.

- This response packet is received at A and is called backscatter on the telescope because it is a reflection to an address that did not initiate the communication.

Telescope operators can then measure internet events by the backscatter they receive. These operators have a window on internet traffic, and certain attacks, that is unique.

Researchers at the Merit Network were unable to find evidence of Muddling Meerkat responses in their backscatter data. The Merit Network researchers subsequently reached out to the operators of another large telescope at the Center for Applied Data Analysis (CAIDA), to see if Muddling Meerkat had spoofed querier IP addresses in the ranges monitored by the CAIDA telescope.[25] CAIDA had no captured backscatter associated with this activity. When we combine their results with the earlier observations of large-scale DNS queries emitting from Chinese IP addresses, we are confident that Muddling Meerkat is not broadly spoofing querier IP addresses in its operations. This is a major difference between Muddling Meerkat and ExploderBot.

THE ROLE OF CHINESE IP ADDRESSES

Because of the complexity involved in Muddling Meerkat operations and the impact of the GFW, it is challenging to determine whether specific events with Chinese IP addresses are "real." What I mean here by "real" is that it can be unclear whether a specific IP address is "answering" a query as a result of the GFW. Similarly, it can be difficult to separate spoofed IP addresses from those that originated queries.

Our approach to this problem was to draw conclusions from overall statistics. As explained earlier, in the section entitled IPv4 Records for Random Subdomains, we observed that Chinese IP addresses "answered" Muddling Meerkat queries where that IP address is known to not have port 53 open. With a large number of these types of examples, we can conclude that "answers" are results of the GFW and not "real" answers.

When we look at querier behavior, some IP addresses stand out. These IP addresses occur with a much higher frequency than the open resolver IPs. They are the source of queries that were outside of the normal resolution for DNS, including to IP addresses that were hosting open resolvers. Some of these querier IP addresses have been repeatedly reported for aggressive scanning and other questionable practices.[26] Table 2 presents an example of source IP addresses and queries.

---

24 https://en.wikipedia.org/wiki/Network_telescope

25 https://www.caida.org/

26 https://www.abuseipdb.com/check/183.136.225.14?page=8

Querier IP Address	Query Name
183[.]136[.]225[.]45	ybzz[.]kb[.]com, xv9k[.]kb[.]com, 0h5w[.]kb[.]com
183[.]136[.]225[.]14	y4fw[.]kb[.]com, mq5i[.]kb[.]com, h420[.]kb[.]com

Table 2. Sample querier IP addresses and queries observed in January 2024. These IP addresses were not hosting open resolvers as of January 31, 2024. Some of these queries were directed at known open resolvers.

LOCATING MUDDLING MEERKAT ACTIVITY

We can observe Muddling Meerkat in part from several sources. Recursive resolvers, like ours, can observe both queries for random subdomains as well as queries for MX records of the target domains. When resolved through the global DNS, the vast majority of these queries will result in an NXDOMAIN response. If there are no open or public resolvers in the network, I don't believe you will see Muddling Meerkat in the DNS logs. Unfortunately, many DNS logging systems record only successful resolutions, and network owners may be blind to the activity because of this limitation.

For those who can observe them, Muddling Meerkat queries are likely to appear intermittently, similar to the examples in Figures 6 and 7, and depend on the size of the network. At Infoblox, we see more Muddling Meerkat traffic than a typical organization would because we resolve DNS queries for customers around the world. Our cloud recursive resolvers handled over 33 trillion queries in 2023 alone.

In addition to DNS query logs, researchers should be able to find traces of Muddling Meerkat in a number of other sources:

- The root, TLD, and authoritative name servers will all have evidence of Muddling Meerkat activity dating back to October 2019 and possibly earlier. Because the actor does not control the target domains, and they are querying broad IP ranges for the records, open resolvers will forward the queries and result in requests at each server within the resolution chain.

- Recursive resolver caches also capture evidence of Muddling Meerkat

- DNS honeypot owners will likely receive queries depending on how broadly Muddling Meerkat queries IP addresses.

- Flow data may contain indications of activity, particularly if it monitors Chinese IP space or shows an unusual variety of port 53 connections to the authoritative name servers, especially arising from open resolver IP addresses.

Queries to any domains provided at the end of this report should be considered suspect. But keep in mind the broad use of these domains for Active Directory and DNS search domains. In addition to the target domain, there should be MX record queries, particularly for short random subdomains. There are other suspect queries for a subset of the Muddling Meerkat domains, which are not included in this report. These are A record queries that appear to leak network information to the authoritative server. However, I am not able to tie this activity definitively to Muddling Meerkat.

ATTRIBUTION AND MOTIVATION

Muddling Meerkat appears to be a Chinese state actor. Because we can observe MX record responses from Chinese IP addresses that are not open on port 53 of Muddling Meerkat target domains over multiple years, I am confident those responses are results of the GFW. At the same time, proper MX responses from the GFW have never been reported before and researchers, including myself, have been unable to trigger the behavior manually. In order to induce selective responses like those we have observed over four years, it seems that Muddling Meerkat must somehow be connected to the GFW operators. While I also don't know how these selective responses are triggered, it is possible that signatures contained in the IP packets, like those observed in ExploderBot traffic, are used to signal a different response from the GFW.

The motivation for these operations is unclear. The data we have suggests that the operations are performed in independent "stages"; some include MX queries for target domains, and others include a broader set of queries for random subdomains. The DNS event data containing MX records from the GFW often occurs on separate dates from those where we see MX queries at open resolvers. Because the domain names are the same across the stages and the queries are consistent across domain names, both over a multi-year period, these stages surely must be related, but we did not draw a conclusion about how they are related or why the actor would use such staged approaches.

Given the research conducted thus far, here are some thoughts on possible motivations:

- Is it a DDoS attack? No, at least not in the current form. The volume of queries observed is far too low to impact authoritative servers or intermediate resolvers. There is no indication there is a reflection attack involved either.

- Is it data exfiltration? This is highly unlikely. The actor does not control the authoritative name servers, uses short subdomain labels with minimal ability to carry information, appears to broadcast packets widely and does not control the return path.

- Is it an open resolver scan? Also unlikely. Of the many ways to find open resolvers, all are simpler than what we observe in these events.

- Is it an internet mapping effort? Well, possibly. Though it seems like a highly convoluted operation to map networks.

- Is it pre-positioning for DDoS attacks? Possibly. To be effective for DDoS, the actor would need to change the operation significantly.

- Is it internet research of some kind? Possibly. If so, it is a very long-running research program and one without a clear aim that I can discern.

- Is it the result of a software bug or some other application? No. This explanation was previously posed by skeptics in response to the ExploderBot research that we conducted. Nothing in the data supports the conclusion that these are incidental DNS queries. Muddling Meerkat activities are very deliberate and very clever.

Is it possible that some other state actor is pretending to be the GFW and spoofing both queries and responses? Many things are possible, not all are plausible.

CONCLUSION AND RECOMMENDATIONS

When you spend as much time as I do staring at DNS, you sometimes wonder if there is anything normal in it. After years of working in this field, I still regularly learn new things and observe new actor behavior. Often, we discover a new actor as a result of some other unrelated factor. In this case, investigating an illegal Chinese gambling network led me to discover anomalous MX records. After chasing a number of red herrings, I formed a clearer picture of the Muddling Meerkat operations when I collaborated with external researchers to share data and analysis. In the end, although I'm writing this report, the analysis and conclusions are the result of joint work where different parties all brought a different perspective to uncover previously undocumented behavior of the GFW and a mysterious multi-year DNS operation.

Our research also highlights potential network vulnerabilities that arise from neglect and the complexity of modern internet communications. In particular, I recommend that network administrators:

- Actively seek out and eliminate open resolvers in their networks. Identifying these devices can be challenging, but companies like Infoblox and organizations like the Shadow Server Foundation can offer critical information to help.

- Do not use domains that you do not own for Active Directory or DNS search domains. You are very likely to leak information about your network and user applications to the authoritative name server, as well as to other appliances outside of your control. This kind of information can allow a bad actor to perform passive reconnaissance of the network for targeted attacks.

- Incorporate DNS detection and response (DNSDR) into your security stack. Only a DNS resolver can effectively handle threats that are inherent in DNS. Most security products won't even recognize the difference between an MX query and an A record query.

- Report Muddling Meerkat activity to the community. Because it is impossible to observe the entire scope from any one vantage point, it is important to crowdsource an understanding of this threat. In particular, reporting additional Muddling Meerkat domains will help others find open resolvers and activity in their network.

Ultimately, I share the concerns expressed by CISA about the PRC and the threat of pre-positioning for cyberattacks globally. In my professional experience, I have found Chinese threat actors to be extremely adept at managing, understanding, and leveraging the DNS for many purposes—whether that be censorship, cybercrime, or DDoS attacks. They also have some of the finest researchers in the field. Whatever the real goal of Muddling Meerkat is, we should not underestimate the talent and patience of the PRC to achieve it.

INDICATORS OF ACTIVITY (TARGET DOMAINS)

Note that these domains are not indicators of compromise or necessarily malicious. Some of the domains used by Muddling Meerkat are parked, others host gambling sites and other possibly illegal content, and others are active legitimate domains. The full scope of Muddling Meerkat target domains is likely much larger.

These domains host no website, host illegal content, or are parked. They likely can be blocked without impact: 4u[.]com, kb[.]com, oao[.]com, od[.]com, boxi[.]com, zc[.]com, s8[.]com, f4[.]com, b6[.]com, p3z[.]com, ob[.]com, eg[.]com, kok[.]com, gogo[.]com, aoa[.]com, gogo[.]com, zbo6[.]com, id[.]com, mv[.]com, nef[.]com, ntl[.]com, tv[.]com, 7ee[.]com, gb[.]com, tunk[.]org, q29[.]org

These domains host websites and blocking them may negatively affect your network: ni[.]com, tt[.]com, pr[.]com, dec[.]com

IP addresses used to launch attacks:
- 183[.]136[.]225[.]45
- 183[.]136[.]225[.]14

INFOBLOX THREAT INTEL

Infoblox Threat Intel is the leading creator of original DNS threat intelligence, distinguishing itself in a sea of aggregators. What sets us apart? Two things: mad DNS skills and unparalleled visibility. DNS is notoriously tricky to interpret and hunt from, but our deep understanding and unique access give us a backstage pass to the internet's inner workings. We're proactive, not just defensive, using our insights to disrupt cybercrime where it begins. We also believe in sharing knowledge to support the broader security community by publishing detailed research and releasing indicators on GitHub. In addition, our intel is seamlessly integrated into our Infoblox DNS Detection and Response solutions, so customers automatically get the benefits of it, along with ridiculously low false positive rates.

Infoblox unites networking and security to deliver unmatched performance and protection. Trusted by Fortune 100 companies and emerging innovators, we provide real-time visibility and control over who and what connects to your network, so your organization runs faster and stops threats earlier.

Corporate Headquarters
2390 Mission College Blvd, Ste. 501
Santa Clara, CA 95054

+1.408.986.4000
www.infoblox.com