Identifier Technology Health Indicators (ITHI)
Metric Collection M3, M4, M6

Christian Huitema (Private Octopus Inc.), Alain Durand (ICANN)

ICANN 60, Abu Dhabi, 28 October - 3 November 2017
In this talk

• Definition of DNS-related ITHI metrics:
  • M3: overhead in root traffic,
  • M4: usage of TLDs and leakage of undelegated strings,
  • M6: usage of IANA-registered DNS parameters.

• Proposed methodology and tools
  • Ask for cooperation from operators of recursive resolvers
M3: overhead in root traffic
ITHI M3: Overhead in Root Traffic

- Overhead at root needs tracking
  - Many “NX Domain” responses
  - Many queries not needed if resolver caches were TTL compliant

- Proposing three metrics:
  - M3.1: NX Domains/Total Queries
  - M3.2: % not TTL compliant queries
  - M3.3: NX Domain per classes of TLD

Example of results, from the analysis of some B-Root traces
ITHI M3.3: NX Domain per classes of TLD

- M3.3.1: RFC 6761 “Special Usage” names, e.g. “.LOCAL”
- M3.3.2: Frequently leaked names, e.g. “.HOME”
- M3.3.3: Suspected automatic generation, e.g. “.FTTPFPTPXGVWJO”
- M3.3.4: all others

Example of results, from the analysis of some B-Root traces
ITHI M3.3.1: Overhead per RFC 6761 Names

• RFC 6761
  • IETF defines “special use” domain names, including some special use TLD
  • Names should never be found in DNS queries, or sent to the root
  • Yet they leak...

• ITHI Metric M3.3.1
  • Track % of overhead for RFC 6761 TLD

<table>
<thead>
<tr>
<th>RFC 6761 TLD</th>
<th>%Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>.LOCAL</td>
<td>...%</td>
</tr>
<tr>
<td>.INVALID</td>
<td>...%</td>
</tr>
<tr>
<td>.LOCALHOST</td>
<td>...%</td>
</tr>
<tr>
<td>.TEST</td>
<td>...%</td>
</tr>
<tr>
<td>.ONION</td>
<td>...%</td>
</tr>
<tr>
<td>.EXAMPLE</td>
<td>...%</td>
</tr>
</tbody>
</table>
ITHI M3.3.2: Overhead by Frequent Names

• M3.3.2:
  • List of most frequently appearing non registered domains

• Methodology
  • Find the “most frequent” non registered domains in traces
  • Retains the names that cause more than 0.1% of leaks

<table>
<thead>
<tr>
<th>TLD</th>
<th>%overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>...%</td>
</tr>
<tr>
<td></td>
<td>...%</td>
</tr>
<tr>
<td></td>
<td>...%</td>
</tr>
<tr>
<td></td>
<td>...%</td>
</tr>
<tr>
<td></td>
<td>...%</td>
</tr>
</tbody>
</table>
ITHI M3.3.3: Overhead by Automatic Names

• Some overhead correspond to suspected automatically generated names

• M3.3.3:
  • Define suspected “patterns” (TBD)
  • Count names that match patterns that account for more than 0.1% of traffic

<table>
<thead>
<tr>
<th>Pattern</th>
<th>%overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>pattern_1</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_2</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_3</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_4</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_5</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_6</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_7</td>
<td>... %</td>
</tr>
<tr>
<td>pattern_8</td>
<td>... %</td>
</tr>
<tr>
<td>...</td>
<td>... %</td>
</tr>
</tbody>
</table>
ITHI M3.3.4: Other Overhead

• Capture a variety of overhead sources, not accounted for by M3.3.1, M3.3.2, M3.3.3

• Defined as difference
  • Total NX Domains = M3.3.1 + M3.3.2 + M3.3.3 + M3.3.4

• Note: M3.3.3 only computed on TLD not found in M3.3.1, M3.3.2
M4: Usage of TLDs and Leakage of Undelegated Strings
M4: Usage of TLDs and Leakage of Undelegated Strings

• M4.1: Usage volume of delegated TLD
  • For each delegated TLD, fraction of queries directed at <TLD>

• M4.2: Leakage of RFC 6761 Special Use Names
  • For each RFC 6761 name, fraction of queries directed at <name>

• M4.3: Leakage of frequent non delegated strings
  • Find most frequent non delegated top level strings in queries
  • Retain name if fraction > 0.1%, List < string>, fraction of query

• M4.4: Leakage of other strings
  • All queries at non registered strings not in M4.2, M4.3
M6: usage of IANA-registered DNS parameters in DNS queries
M6: usage of IANA-registered DNS parameters in DNS queries

- IETF
  - Protocol Definition

- IANA
  - Parameter registry

- Dev.
  - Registration
  - (or squatting)

- ?
  - Practical usage

- Example of registries
  - DNS RR Types
  - EDNS OPT Types
  - DNSSEC Algorithms

- Two questions
  - Are the registered values used in DNS queries?
  - Do we observe squatting?
Metric Definition, Fictitious Example, Registry with 16 possible entries

\[
\text{Usage} = \frac{\text{Nb used}}{\text{Nb registered}} = \frac{7}{10} = 70\%
\]

\[
\text{Squatting} = \frac{\text{Volume Non Reg.}}{\text{Volume Total}} = \frac{8}{68} = 11.8\%
\]
M6.X.N.1, 2 and 3

• Multiple registries
• Registry Index, form X.N
  • X: one of DNS, DANE, DNSSEC
  • N: index of specific registry in the group specified above
• Three metrics per registry
  • M6.X.N.1: Usage
  • M6.X.N.2: Squatting
  • M6.X.N.3.V: Volume, for each registered value “V”

• Example: RR Type
  • DNS Registry number 2
  • M6.DNS.2.1: usage metric for RR Types
  • M6.DNS.2.2: squatting metric for RR Types
  • M6.DNS.2.3.28: usage of value 28 (AAAA)
<table>
<thead>
<tr>
<th>Group</th>
<th>Parameters</th>
<th>Metric Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>DANE</td>
<td>TLSA Certificate Usages</td>
<td>M6.DANE.1</td>
</tr>
<tr>
<td></td>
<td>TLSA Selectors</td>
<td>M6.DANE.2</td>
</tr>
<tr>
<td></td>
<td>TLSA Matching Types</td>
<td>M6.DANE.3</td>
</tr>
<tr>
<td>DNS</td>
<td>DNS CLASSes</td>
<td>M6.DNS.1</td>
</tr>
<tr>
<td></td>
<td>Resource Record (RR) TYPES</td>
<td>M6.DNS.2</td>
</tr>
<tr>
<td></td>
<td>DNS OpCodes</td>
<td>M6.DNS.3</td>
</tr>
<tr>
<td></td>
<td>DNS RCODEs</td>
<td>M6.DNS.4</td>
</tr>
<tr>
<td></td>
<td>AFSDB RR Subtype</td>
<td>M6.DNS.5</td>
</tr>
<tr>
<td></td>
<td>DHCID RR Identifier Type Codes</td>
<td>M6.DNS.6</td>
</tr>
<tr>
<td></td>
<td>DNS Label Types</td>
<td>M6.DNS.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Parameters</th>
<th>Metric Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS SEC</td>
<td>DNS Security Algorithm Numbers</td>
<td>M6.DNSSEC.1</td>
</tr>
<tr>
<td></td>
<td>DNS KEY Record Diffie-Hellman Prime Lengths</td>
<td>M6.DNSSEC.2</td>
</tr>
<tr>
<td></td>
<td>DNS KEY Record Diffie-Hellman Well-Known Pairs</td>
<td>M6.DNSSEC.3</td>
</tr>
<tr>
<td>DNS</td>
<td>DNS EDNS0 Option Codes (OPT)</td>
<td>M6.DNS.8</td>
</tr>
<tr>
<td></td>
<td>DNS Header Flags</td>
<td>M6.DNS.9</td>
</tr>
<tr>
<td></td>
<td>EDNS Header Flags (16 bits)</td>
<td>M6.DNS.10</td>
</tr>
<tr>
<td></td>
<td>EDNS version Number (8 bits)</td>
<td>M6.DNS.11</td>
</tr>
<tr>
<td></td>
<td>Child Synchronization (CSYNC) Flags</td>
<td>M6.DNS.12</td>
</tr>
</tbody>
</table>
Capture and Computation
Proposed Methodology

• Process:
  • Use multiple collection points
  • At each collection point, collect about 1 million transactions
  • ICANN receives summary data once a day from collection points
  • ICANN aggregate summaries to compute the metrics.

• Open source collection tool provided by ICANN:
  • Removes PII information from the observed data.
  • Produces summary table at collection point.
  • Computes metrics after aggregation
At ICANN:

Network share, e.g. Dropbox:

Monitoring tool (e.g. DNSCAP) → Capture (PCAP) → ITHITOOLS extraction

Summary (CSV)

At specific location, 1M queries at random time of day:

The tool chain for ITHI metrics M3, M4 and M6

ITHITOOLS summary

Merged Summary (CSV) → Metrics (CSV)
Difference between M4 and M3.3

- M3.3 measures overhead at the root
- M4 measures usage and leakage at recursive resolver
- With “perfect” resolvers, M3.3 tends towards 0%, due to caching
  - E.g., NSEC3 aggressive
- M6 mostly observable at resolvers
  - Caching, QName minimization
Dealing With Privacy Issues

- DNS traffic is privacy sensitive
  - IP addresses of users
  - Domain names of servers
  - Patterns of user queries
- We do not need PII data for M3, M4 and M6
  - No need for source IP addresses, queried names
  - Just statistics, no GDPR issue
- Solution: produce aggregated summaries
  - Typical summary size: 8 to 16 KB
ITHI Tool Design

- ITHITOOLS: single tool, three functions:
  - Parse a capture file, produce a summary
  - Merge several summaries
  - Compute the metrics

- Open source:
  - https://github.com/private-octopus/ithitools
  - MIT license
  - C++, Can be built on Windows and Linux

- Can run in a “sand box”
  - No network access required,
  - Summaries can be copied to network share by script
Summary
DNS Recursive Resolvers
Operators, we need your help!

- ITHI metrics help the whole community
  - M3: health of the DNS root
  - M4: analysis of TLD usage and leakage of strings
  - M6: health of IANA parameter registries for DNS

- Capture methodology is safe
  - Minimal load, no privacy issues

- Please contact us if you are interested!