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

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

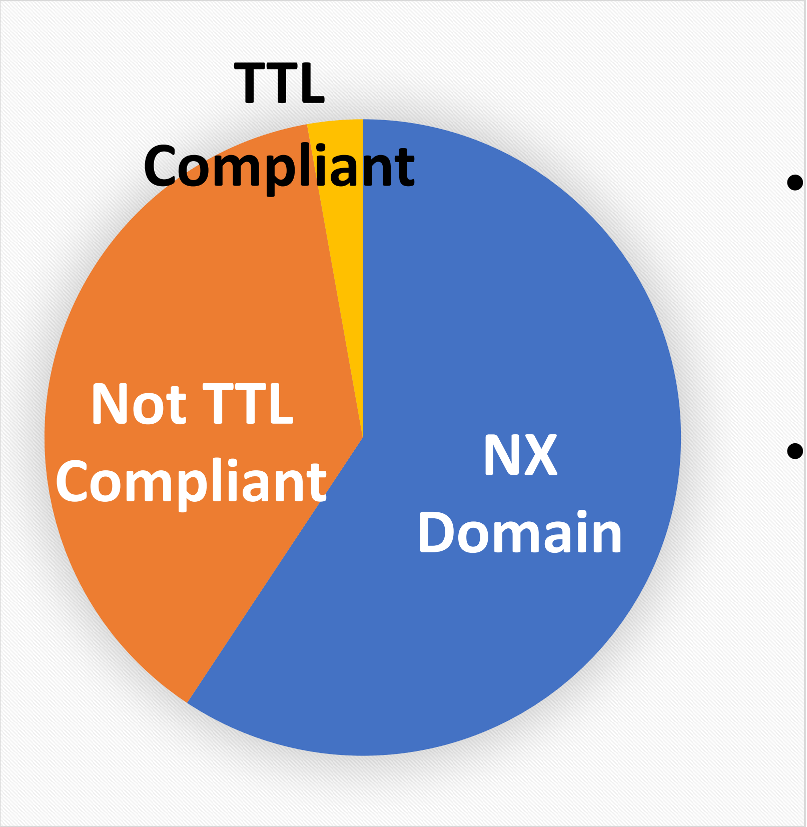
ICANN 60, Abu Dhabi, 28 October - 3 November 2017

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

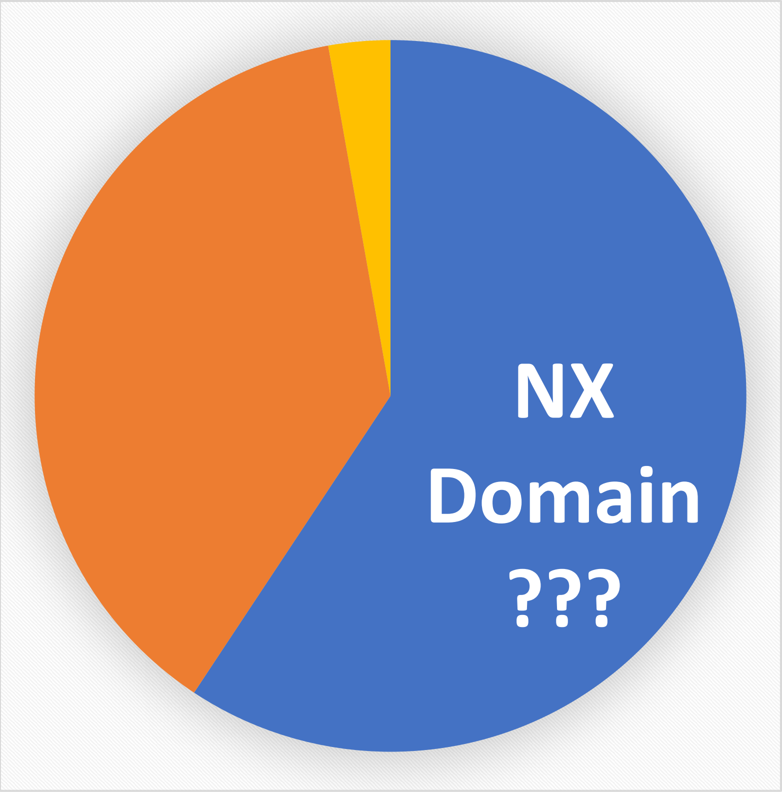
ITHI M3: Overhead in Root Traffic



Example of results, from the analysis of some B-Root traces

- 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

ITHI M3.3: NX Domain per classes of TLD



Example of results, from the analysis of some B-Root traces

- 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. “.FTTPFTPXGVWJO”
- M3.3.4: all others

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

RFC 6761 TLD	%Over-head
.LOCAL	…%
.INVALID	…%
.LOCALHOST	…%
.TEST	…%
.ONION	…%
.EXAMPLE	…%

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

TLD	%overhead
…	…%
…	…%
…	…%
…	…%
…	…%

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

Pattern	%overhead
pattern_1	… %
pattern_2	… %
pattern_3	… %
pattern_4	… %
pattern_5	… %
pattern_6	… %
pattern_7	… %
pattern_8	… %
…	… %

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

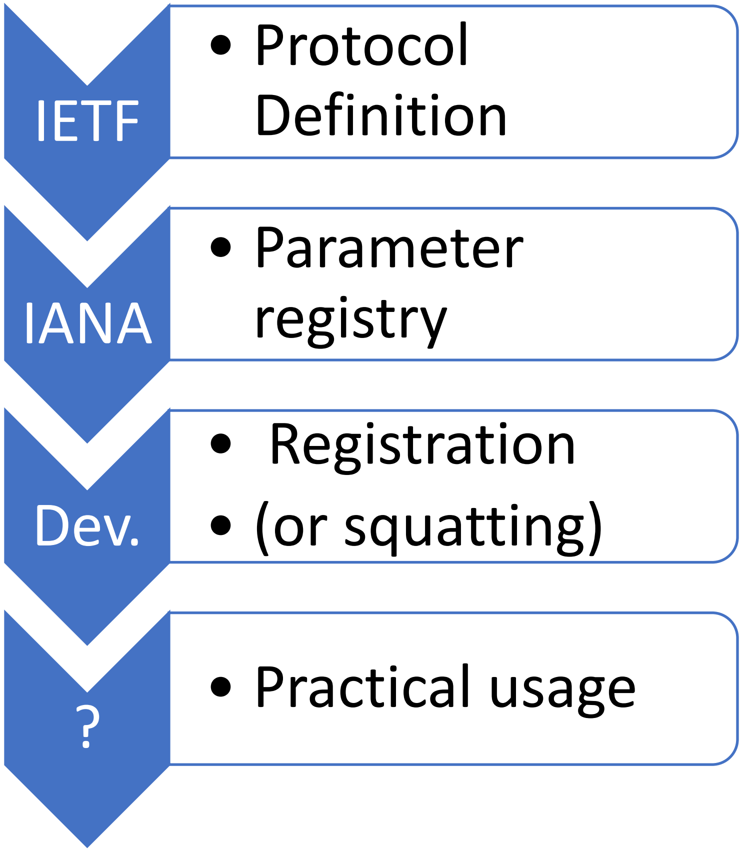
M4: Usage of TLDs and Leakage of Undelegated Strings

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.

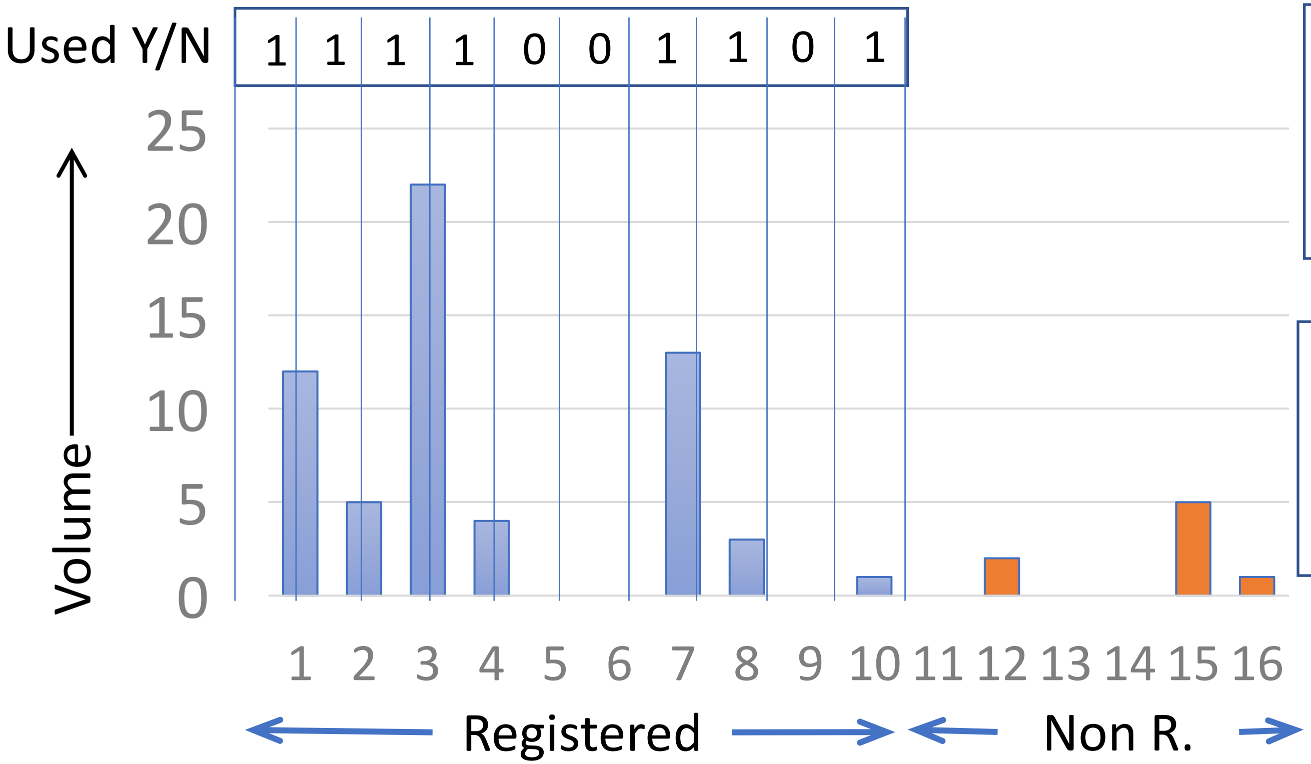
- Registration
- (or squatting)

?

- 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



Used Y/N	1	1	1	1	0	0	1	1	0	1

$$\frac{\text{Nb used}}{\text{Nb registered}} = \frac{7}{10}$$

Usage = 70%

$$\frac{\text{Volume Non Reg.}}{\text{Volume Total}} = \frac{8}{68}$$

Squatting = 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)

List of DNS Parameter Registries

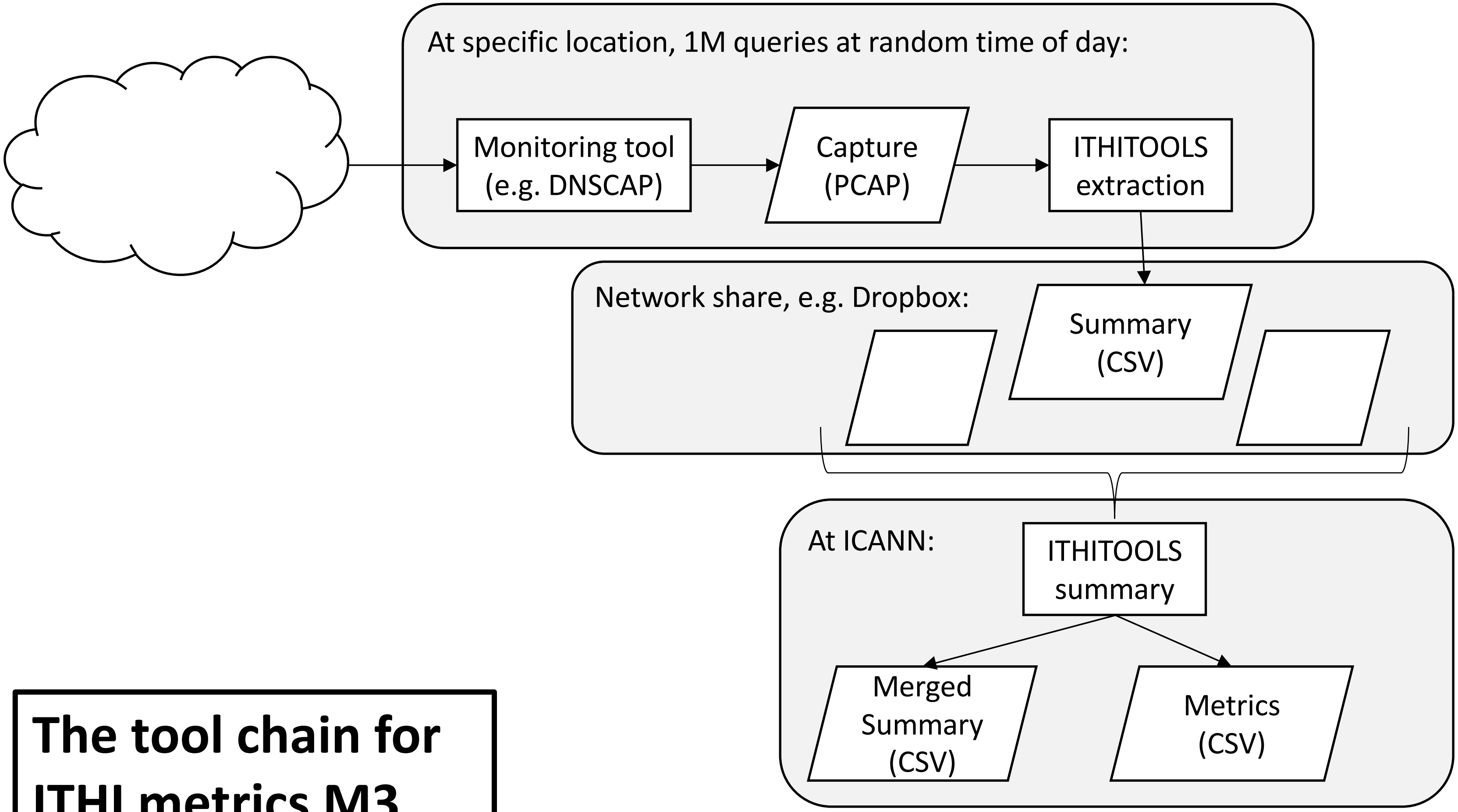
Group	Parameters	Metric Index
DANE	TLSA Certificate Usages	M6.DANE.1
	TLSA Selectors	M6.DANE.2
	TLSA Matching Types	M6.DANE.3
DNS	DNS CLASSes	M6.DNS.1
	Resource Record (RR) TYPEs	M6.DNS.2
	DNS OpCodes	M6.DNS.3
	DNS RCODEs	M6.DNS.4
	AFSDB RR Subtype	M6.DNS.5
	DHCID RR Identifier Type Codes	M6.DNS.6
	DNS Label Types	M6.DNS.7

Group	Parameters	Metric Index
DNS	DNS EDNS0 Option Codes (OPT)	M6.DNS.8
	DNS Header Flags	M6.DNS.9
	EDNS Header Flags (16 bits)	M6.DNS.10
	EDNS version Number (8 bits)	M6.DNS.11
	Child Synchronization (CSYNC) Flags	M6.DNS.12
DNS SEC	DNS Security Algorithm Numbers	M6.DNSSEC.1
	DNS KEY Record Diffie-Hellman Prime Lengths	M6.DNSSEC.2
	DNS KEY Record Diffie-Hellman Well-Known Prime/Generator Pairs	M6.DNSSEC.3

Capture and Computation

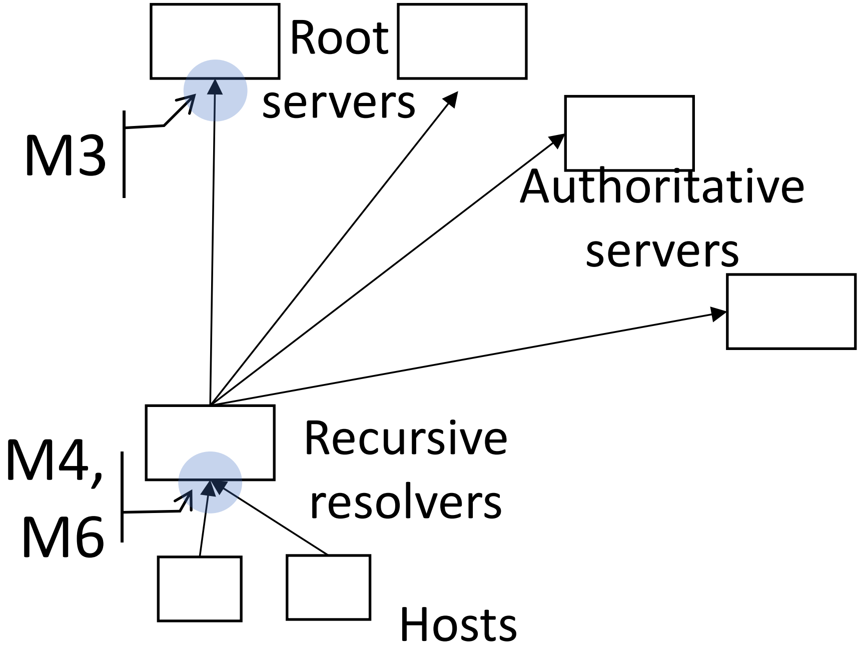
Proposed Methodology

- 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 specific location, 1M queries at random time of day:
Monitoring tool (e.g. DNSCAP)
Capture (PCAP)
ITHITOOLS extraction
Network share, e.g. Dropbox:
Summary (CSV)
At ICANN:
ITHITOOLS summary
Merged Summary (CSV)
Metrics (CSV)

The tool chain for ITHI metrics M3, M4 and M6

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!