ICANN Name Collision Reports Root Cause Analysis

Casey Deccio
NCAP Technical Investigator

Background

- In summer 2021 I was given the responsibility of investigating the name collisions reports submitted to ICANN between 2014 and 2021.
- Initial task: contact each submitter, and find out more details about their submission.
- However, I was not granted permission to contact the submitters.
- Root cause analysis became an exercise in measurement, data collection, and analysis.

Major Questions:

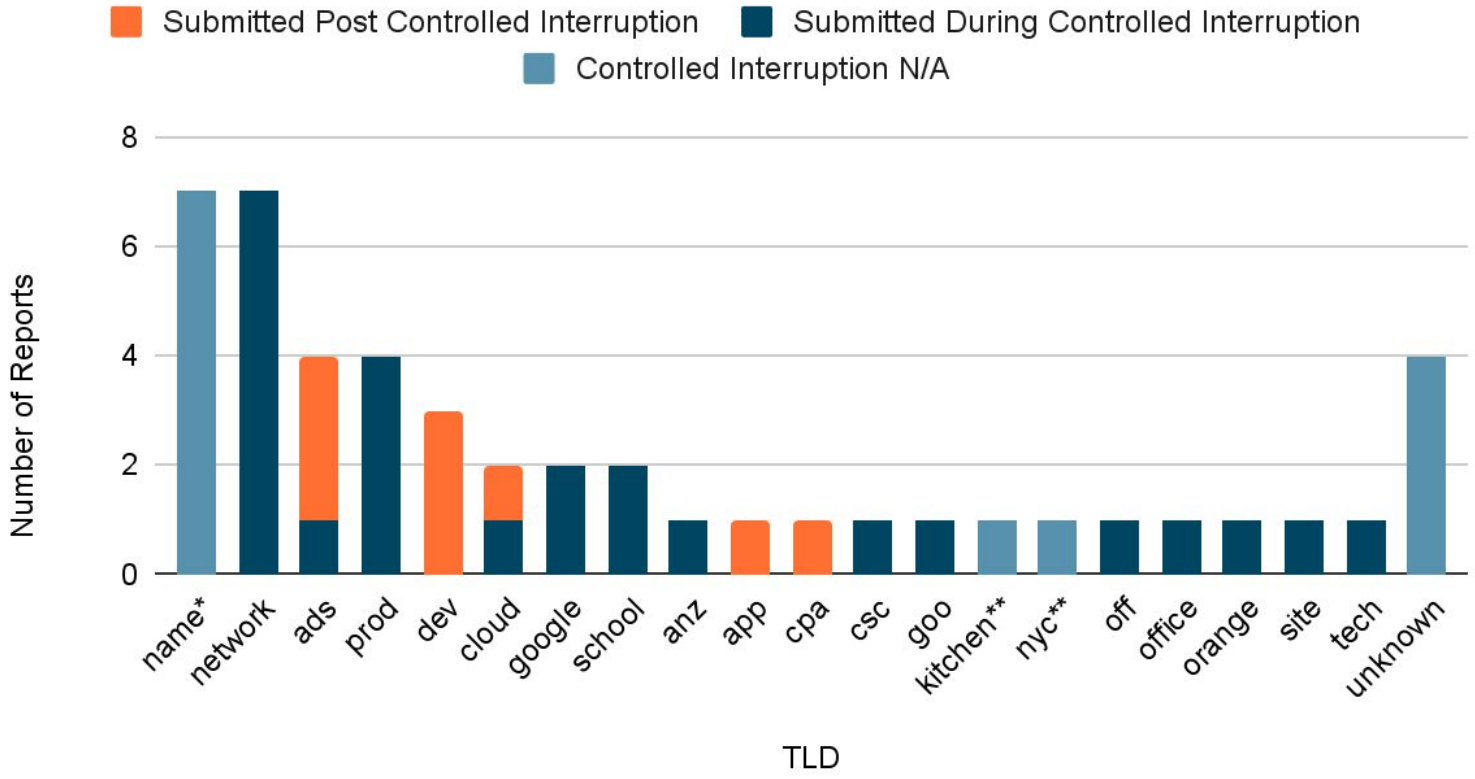
1. What can we learn from the name collisions reports submitted to ICANN?
2. What name collisions were experienced more generally?
3. What was the user/administrator experience with name collisions?

Question 1: What can we learn from the name collisions reports submitted to ICANN?

Name Collision Reports - Overview

- 47 Reports (43 reports include TLD)
  - 7 reports - related to `wpad.domain.name` vulnerability (see other report)
  - 2 reports - new TLDs delegated prior to controlled interruption (`kitchen` and `nyc`)
  - **34 reports - new TLDs delegated after controlled interruption**
    - 25 reports - reported during controlled interruption
    - 9 reports - reported after controlled interruption
- 20 TLDs reported
  - 1 TLD - related to `wpad.domain.name` vulnerability (see other report)
  - 2 TLDs - delegated prior to controlled interruption (`kitchen` and `nyc`)
  - **17 TLDs - delegated after controlled interruption**

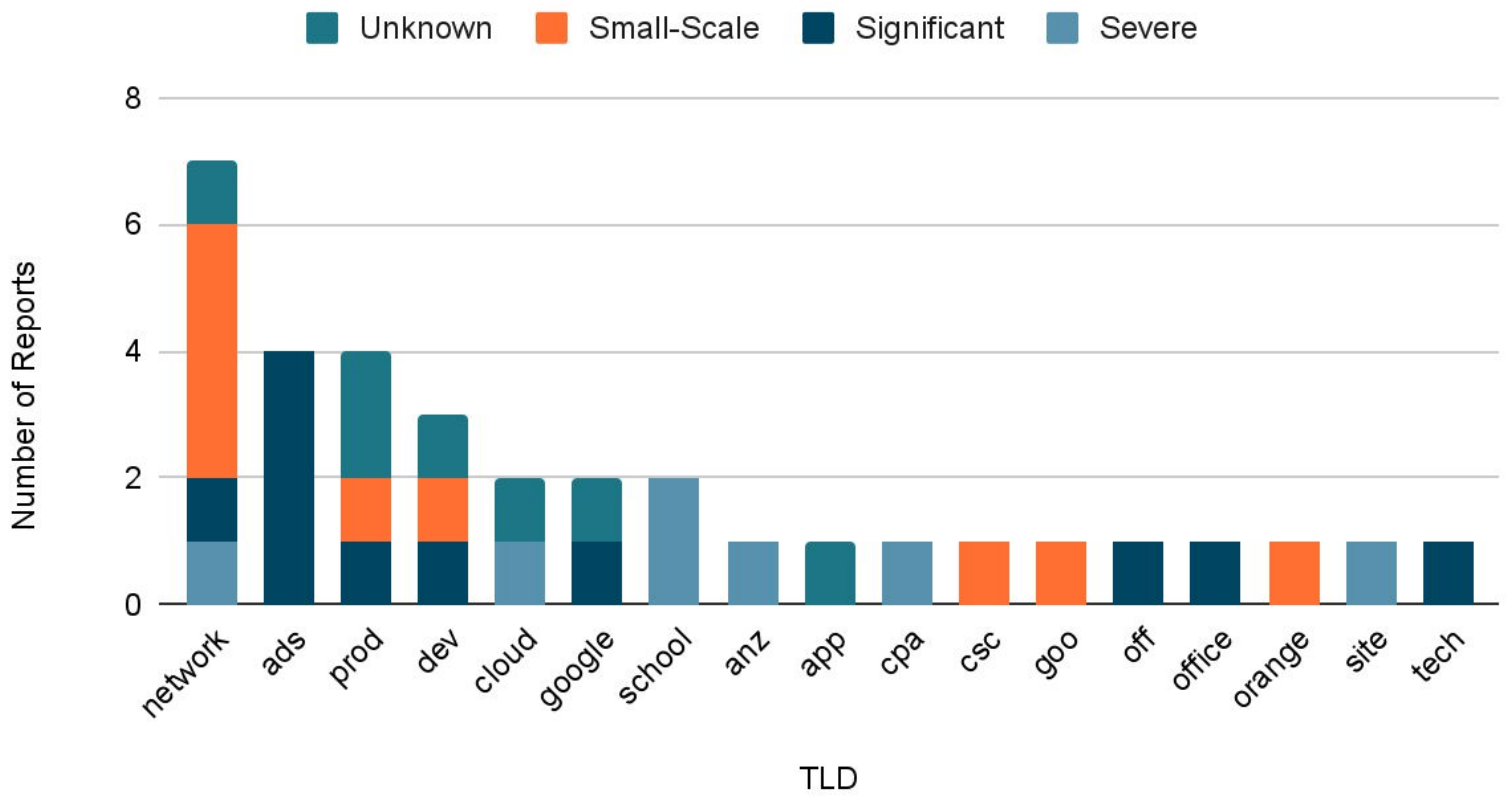
Name Collision Reports - By Submission Date



Name Collision Reports - Severity

- Parties invited to submit report if experiencing “demonstrably severe harm.”
- Reports independently classified (subjectively) by description entered:
  - **Severe: 7**
    - “more 30,000 employees in over 7 countries”,
    - “all of our staff laptops … crash”
  - **Significant: 10**
    - “CRM, MAIL and other Services … do not work correctly”
    - “Unable to resolve internal Hostnames”
  - **Small-Scale: 10**
    - “can't access to some servers”
    - “home network disruption”
  - **Unknown: 7**

Name Collision Reports - Severity



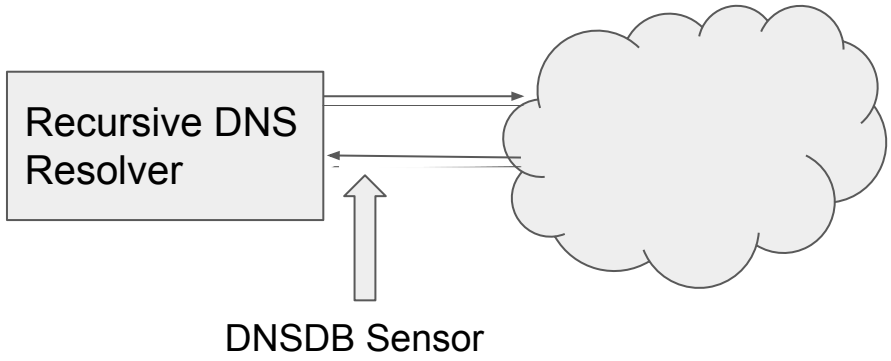
Name Collision Reports - Other Observations

- 127.0.53.53 is only mentioned by 8 (24%) of 34 reports.
- VPN usage is mentioned by 8 reports (33% of the 24 submitted by orgs).
- AD usage is mentioned by 8 reports (33% of the 24 submitted by orgs).

Question 2: What name collisions were experienced more generally?

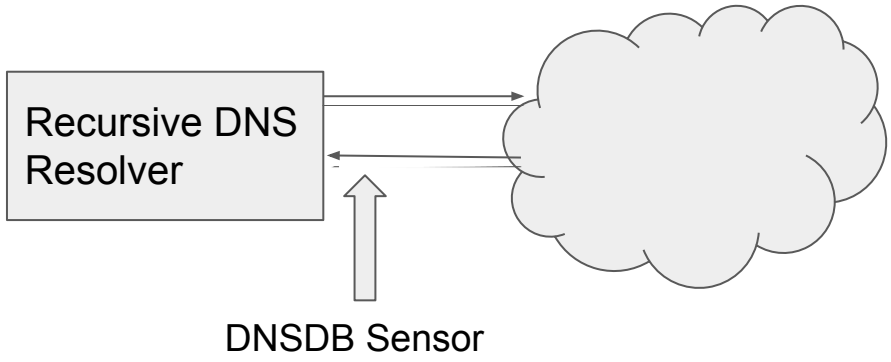
Data Source: DNSDB (by DomainTools)

- DNSDB contains historical DNS name-to-resource mappings.
- Mappings come from DNS responses made at deployed sensors.
- Only positive responses included in DNSDB (i.e., not NXDOMAIN).
- During controlled interruption period for a TLD, *all* responses are positive.
- No IP address available; only query count.



Data Set: Controlled Interruption Queries

- 885 gTLDs delegated between August 2014 (start of controlled interruption) and June 2021.
- Retrieved *every* DNS mapping observed during controlled interruption period for every new gTLD.
- Effective result: every qname/count queried for yet-to-be-delegated TLDs.



Quantifying Name Collisions - Possible Metrics

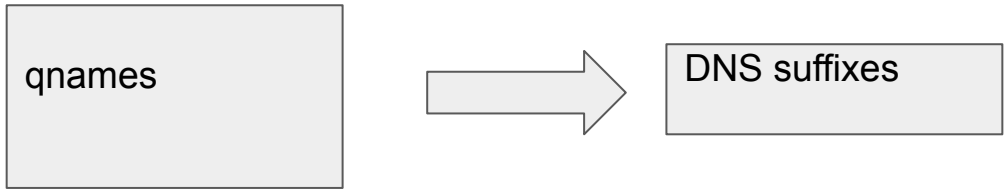
- qname composition:
  - Number of unique qnames - too fine-grained by itself
  - Number of unique SLDs - does not necessarily align with organization or configuration
    - Example: `foo.bar1.baz.com` and `foo.bar2.baz.com`
    - Example: `state.ut.us` and `k12.ut.us`
- Query origin (unavailable with DNSDB):
  - Client IP address count
  - Origin AS count
- Query count:
  - Useful in conjunction with query origin and qname composition

Quantifying Name Collisions - DNS Suffixes

- DNS Suffix
  - Known as “Search domain” (Windows) or “domain” or “search” `resolv.conf` entry (UNIX/Linux).
  - Typically configured by the “network”, either dynamically (e.g., via DHCP) or statically.
  - Used for various purposes:
    - Search list processing for unqualified domains `foo`  `foo.example.com`
    - Web Proxy Auto-Detect (WPAD) `wpad.example.com`
    - ISATAP (IPv6 tunnel gateway detection) `isatap.example.com`
    - Chrome “NXDOMAIN probing” `abdef.example.com` `ghijk.example.com` `lmnop.example.com`

Quantifying Name Collisions - Leaked DNS Suffixes

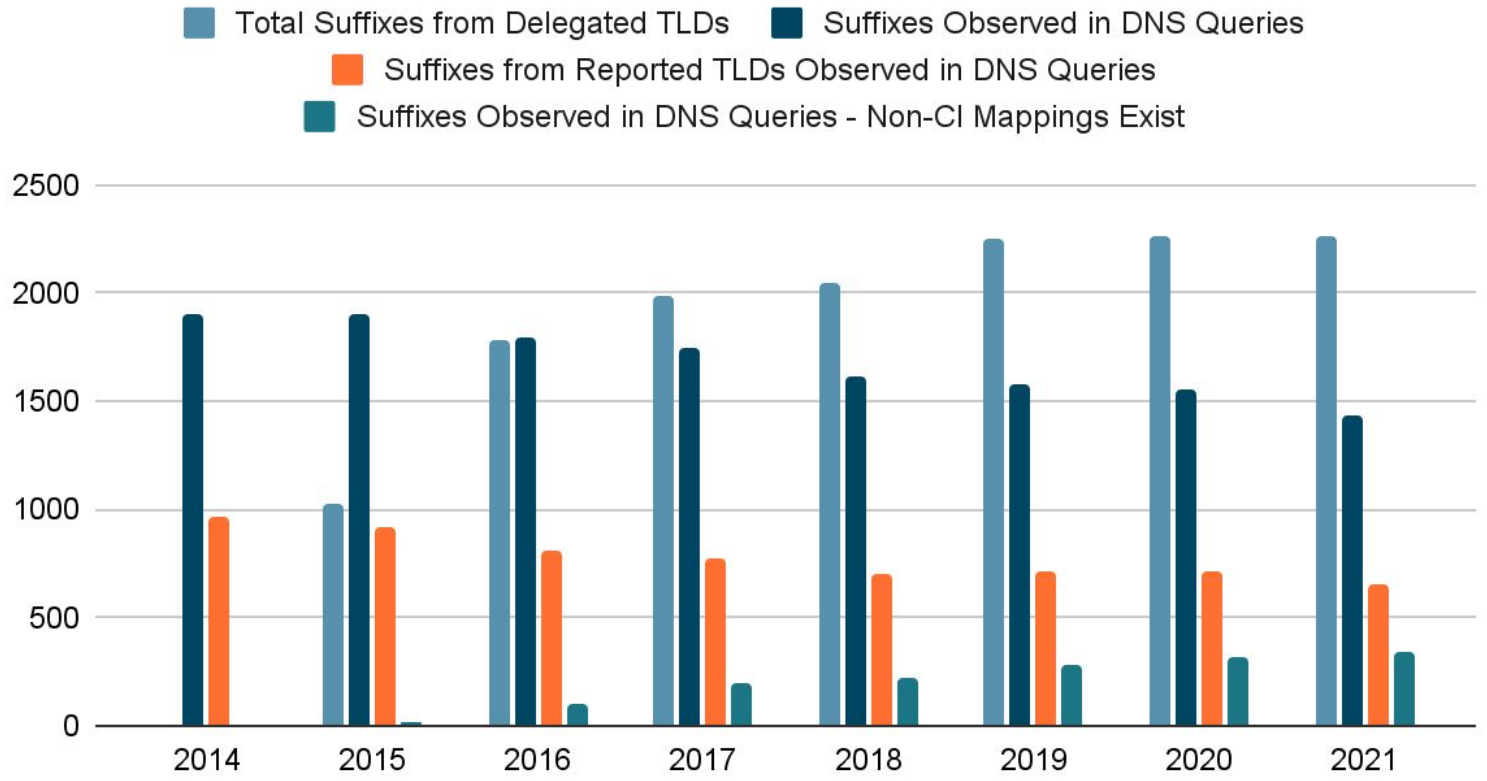
- Extracted DNS suffixes from qnames in DNSDB data using three methods:
  - Inferred Chrome NXDOMAIN probe: 3 one-time queries in 1 second, all with same suffix
  - WPAD DNS query: query observed with `wpad` as first label
  - ISAPTAP DNS query: query observed with `isatap` as first label
- DNS Suffixes extracted: 2,762
  - Includes suffixes from 498 TLDs
- DNS Suffixes reduced to 2,266
  - Excludes TLDs and suffixes from TLDs with low overall suffix counts
  - Includes suffixes from 266 TLDs
  - These become the basis for subsequent analysis



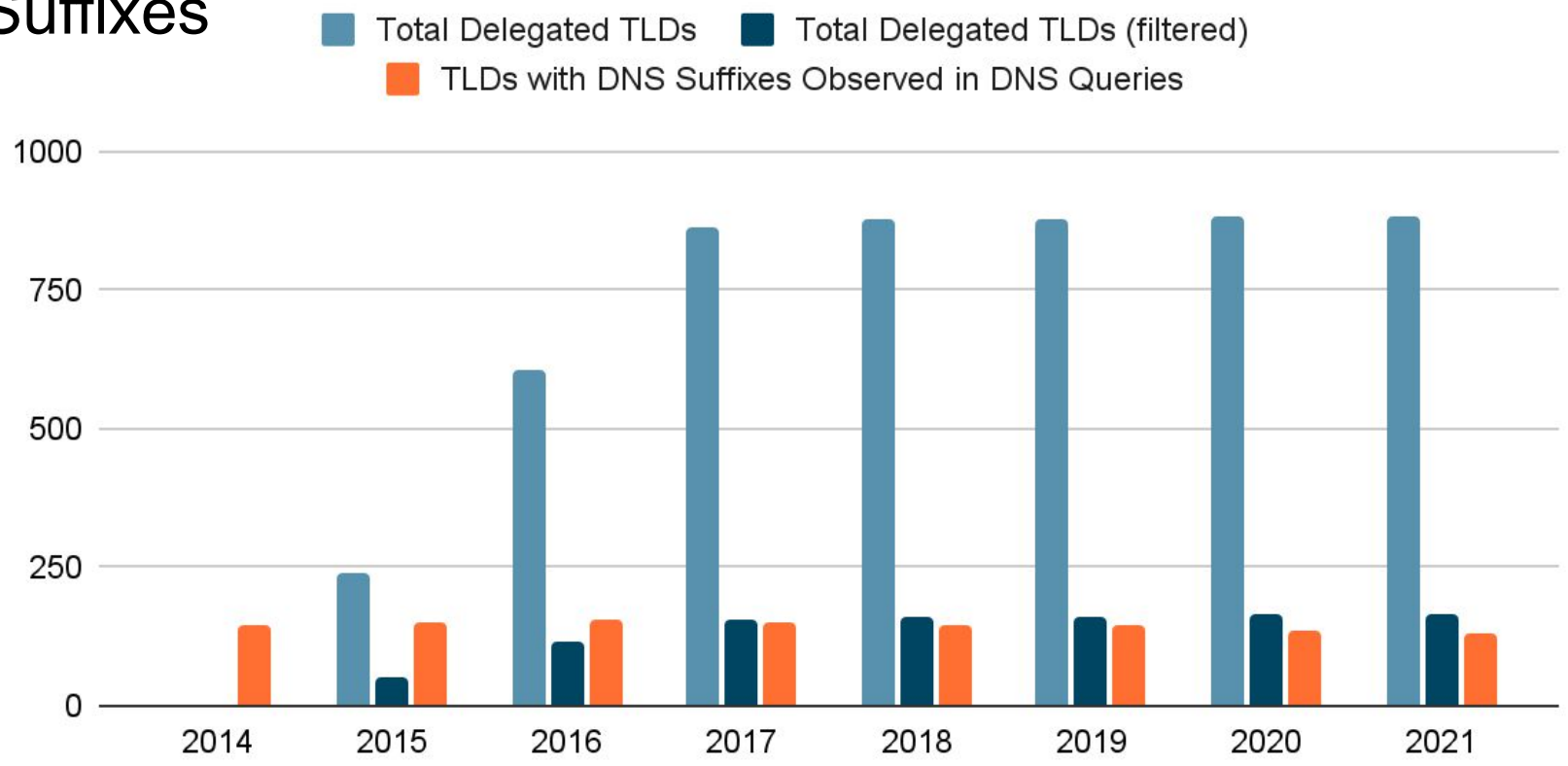
Quantifying Name Collisions - Leaked DNS Queries at Root

- Filtered DNS queries seen at DNS root servers by identified DNS suffixes
- Root servers: A, C, H, and J
- Years: 2014 through 2021

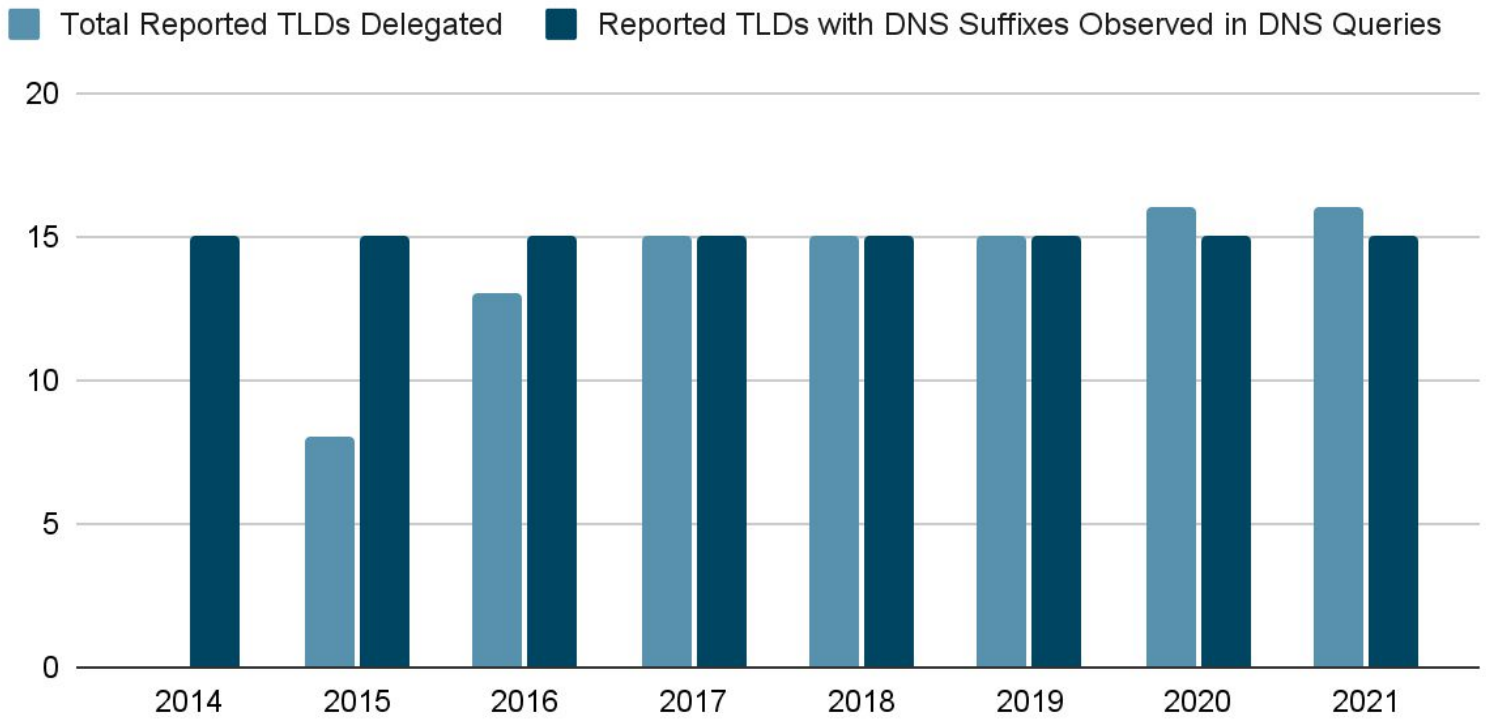
Quantifying Name Collisions - Observed DNS Suffixes



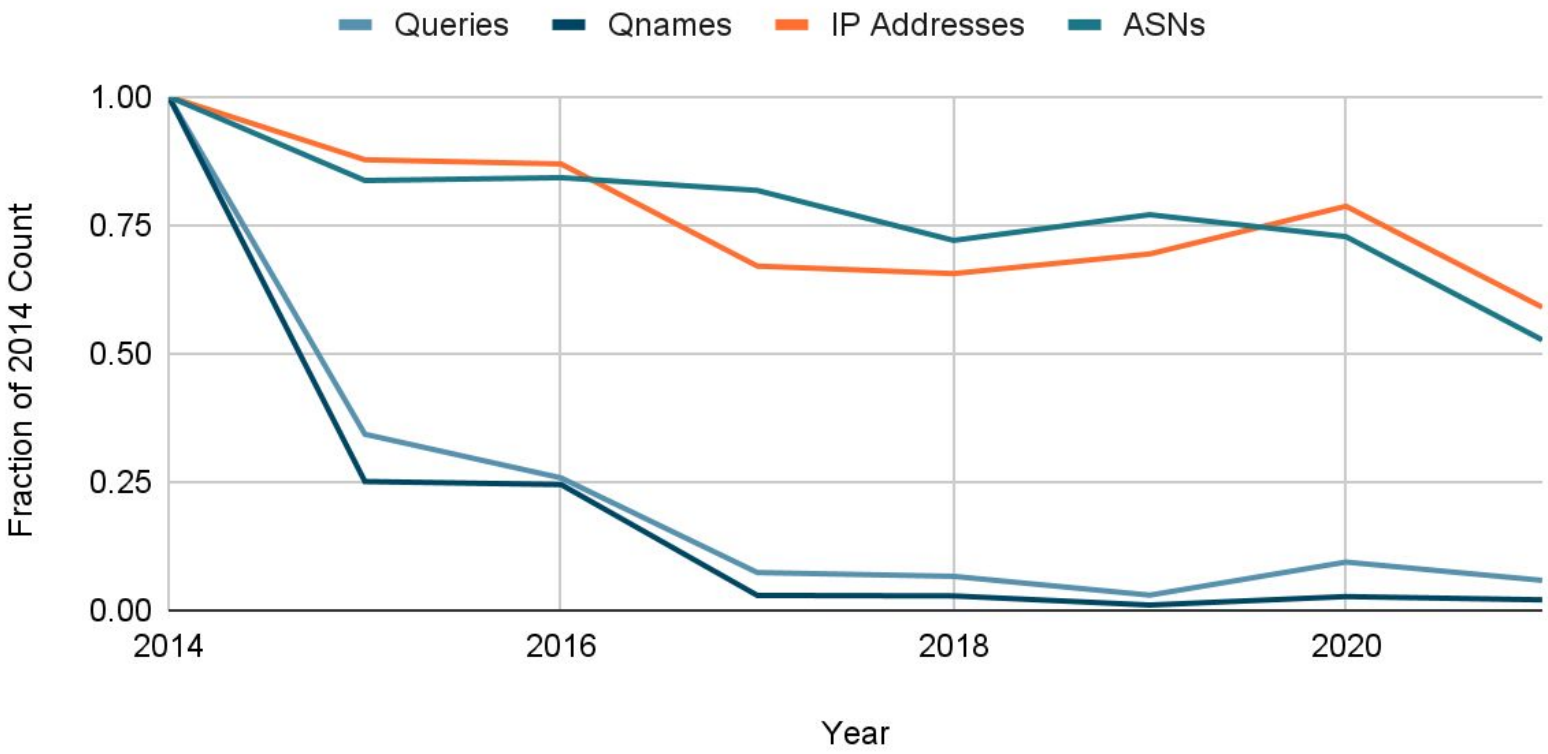
Quantifying Name Collisions - TLDs of Observed DNS Suffixes



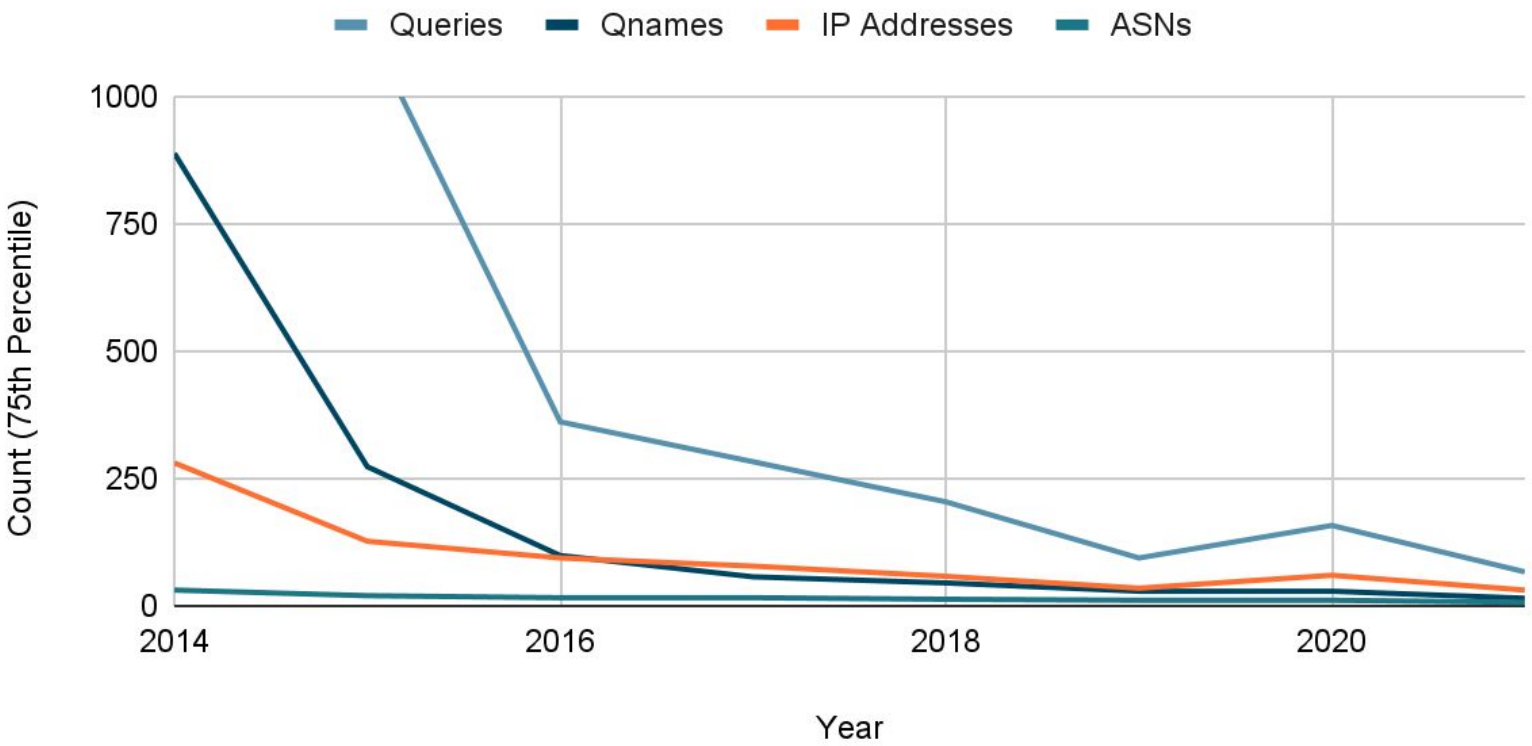
Quantifying Name Collisions - TLDs of Observed DNS Suffixes (Reported TLDs Only)



Quantifying Name Collisions - Overall DNS Queries



Quantifying Name Collisions - Per-Suffix 75th Percentile



Question 3: What was the user/administrator experience with name collisions?

Measuring Impact - Name Collision Report Challenges

- Challenges with ICANN name collisions reports
  - **Bias:** only includes experiences for which:
    - Problems were experienced.
    - Those experiencing problems identified ICANN as the entity to which collisions should be submitted.
    - Presumably, problems experienced resulted in “demonstrably severe harm.”
  - **Result:** no way to reliably measure the following:
    - Those using publicly delegated TLDs as private namespace, experiencing no problems
    - Those that experienced problems but didn’t report them
    - Those that experienced a spectrum of severity

Measuring Impact - Survey on DNS Suffix Usage

- Survey questions
  - Are DNS suffixes under new gTLDs in “private” use by organizations?
  - Which suffixes and TLDs are used?
  - Were problems experienced?
  - Was 127.0.53.53 observed?
  - What was the impact on users and systems?
- Survey distribution
  - **General Survey:** sent to NANOG mail list
  - **Targeted survey:** sent to AS contacts from which leaked private DNS queries were observed
    - Matched DNS suffix to AS description
    - 28 contacts

Measuring Impact - Survey Results

- 10 respondents indicated that their organization used private DNS suffixes.
- 7 respondents indicated problems related to name collisions.
- Problem discovery took days (43%), weeks (14%), or months (43%).
- Problem resolution took days (29%) or years (29%), some unresolved (29%).
- Only 14% of cases indicated that 127.0.53.53 was observed and helpful.
- In 71% of cases, 127.0.53.53 was not observed at all.

Findings

- Private use of DNS suffixes is widespread.
- Name collision reports are supported strongly by measured data.
- Usage of known, private DNS suffixes has decreased over time.
- Controlled interruption is effective at disruption but not at root cause identification.
- Configuring DNS resolvers as authoritative for DNS suffixes is not a panacea.
- The impact of TLD delegation ranged from no impact to severe impact.

Future Work - Identifying “Who” is Impacted and “How Much”

- General observations from analysis:
  - Even statically configured systems are mobile.
  - DNS queries might never leak from their origin ASN.
  - Many ASNs are ISPs.
  - Generic suffixes are in use.
  - Regional subdomain suffixes are in use.
  - Some TLDs are commonly used for Active Directory services.
- Proposal:
  - Automated AS-suffix association.
  - Large-scale reach-out to affected parties.