ENUM and confidentiality

Florian Weimer <fweimer@bfk.de>
BFK edv-consulting GmbH

2nd OARC Workshop, November 2006

Outline

Introduction to ENUM

Risks inherited from DNS

Zone enumeration

DNS cache snooping

Outline

Phone numbers and Internet resources

- Success of end-to-end VoIP seemed to require:
  - integration of traditional touch-tone phones,
  - support of existing processes based on phone numbers,
  - rerouting of PSTN calls over IP networks.
- A service that maps phone numbers to URIs can
  - enable touch-tone phones to address IP services,
  - add new meanings to phone numbers (for example, RFC 2822 mailboxes),
  - routing information (in the form of SIP or H.323 URIs).

The ENUM mapping service

- Basic algorithm is as follows:
  - Reverse the phone number.
  - Split it into digits (which become individual DNS labels).
  - Add some suffix to form a full domain name.
  - Use DNS to map this name to NAPTR RRs.
  - The NAPTR RRs contain regular expressions, rewriting the phone number into URIs.
- `e164.arpa` is the root of the official ENUM tree.
- Example:
  +49721962011 → `1.1.0.2.6.9.1.2.7.9.4.e164.arpa`

The ENUM mapping service (2)

- ENUM fits quite nicely into the DNS architecture.
- DNS provides a highly available distributed database at very low cost.
- Delegation is possible
  - at the country-code level,
  - for individual area codes,
  - for user-assigned numbers and number blocks.
- Incremental dialing is possible in principle, but high query rate is feared (and prevented by typical regexp usage).

BFK

ENUM on DNS: the downsides

- DNS scores highly on availability, but poorly on integrity and confidentiality.
- Telephony services have particularly high integrity and confidentiality requirements.
  - Eavesdropping on phone calls is clearly illegal in most countries.
  - Call data records (who calls who and when?) are strongly protected in many countries.
  - Even the subscriber list needs to be kept private in some jurisdictions.
- Integrity is addressed by DNSSEC (currently at the cost of confidentiality of the subscriber list).
- Confidentiality?

ENUM on DNS: no opt-out

- ENUM-enabled callers will consult DNS
  - even if you have not registered an ENUM domain for your phone numbers,
  - even if your country code has not been delegated from `e164.arpa`.
- This means that some of your call records have already been exposed to the domain name system.
- To some degree, you need to implement this technology to prevent its use.

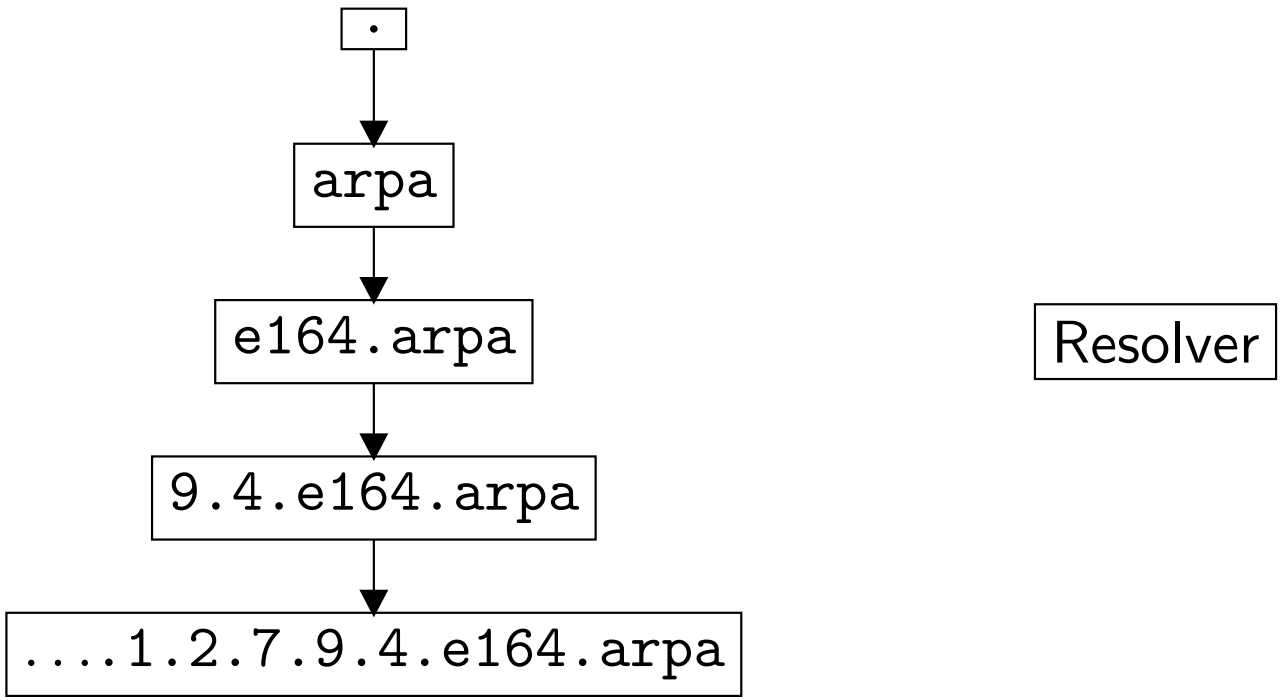
Outline

DNS as a risk

- protocol weaknesses (16 bit message ID, no cryptography)
- protocol complexity (which servers are authoritative for what data?)
- implementation quality
  - the usual suspects, such as buffer overflows
  - errors due to protocol complexity and ambiguity
- However, DNS is apparently good enough to power a multi-billion-dollar industry.

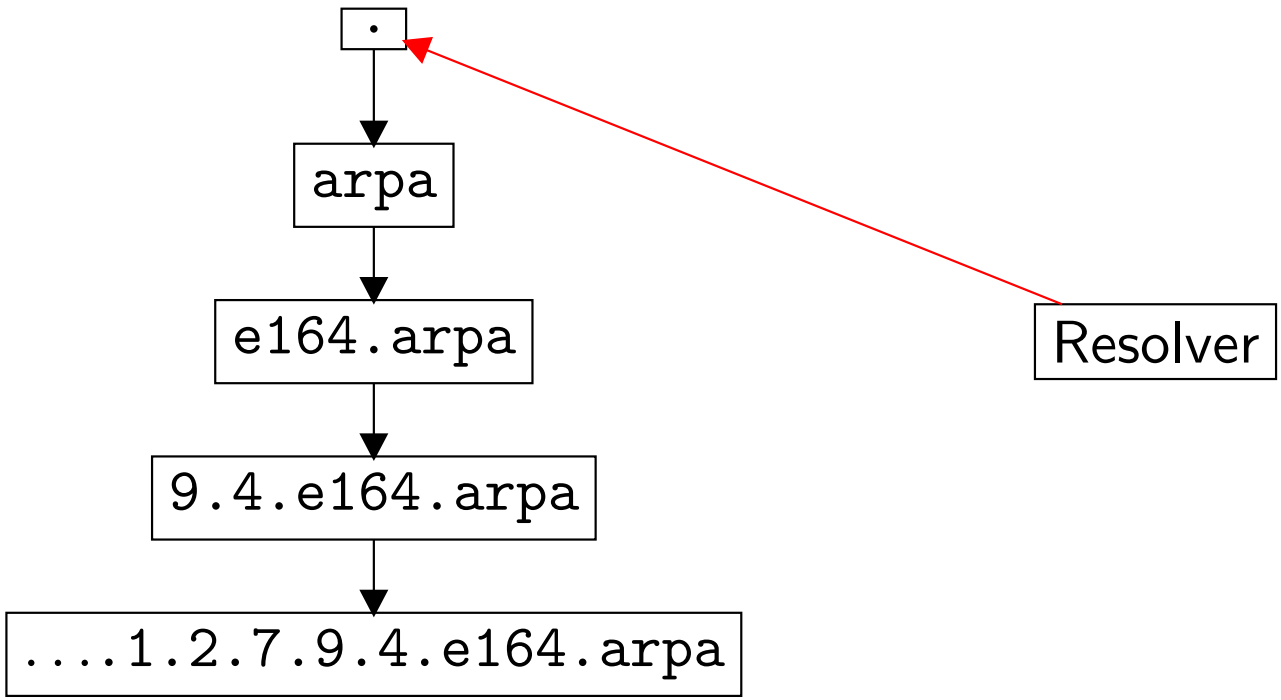
Data leaks along the DNS delegation tree

- Cold cache behavior



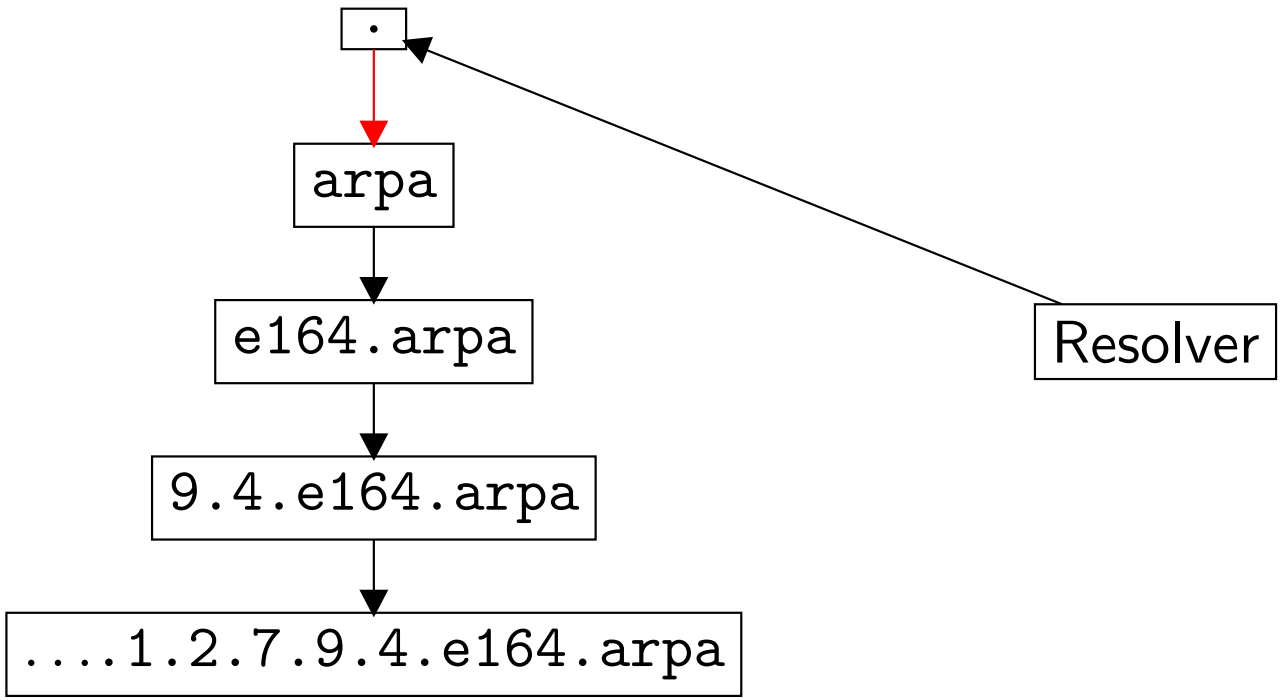
Data leaks along the DNS delegation tree

- Cold cache behavior



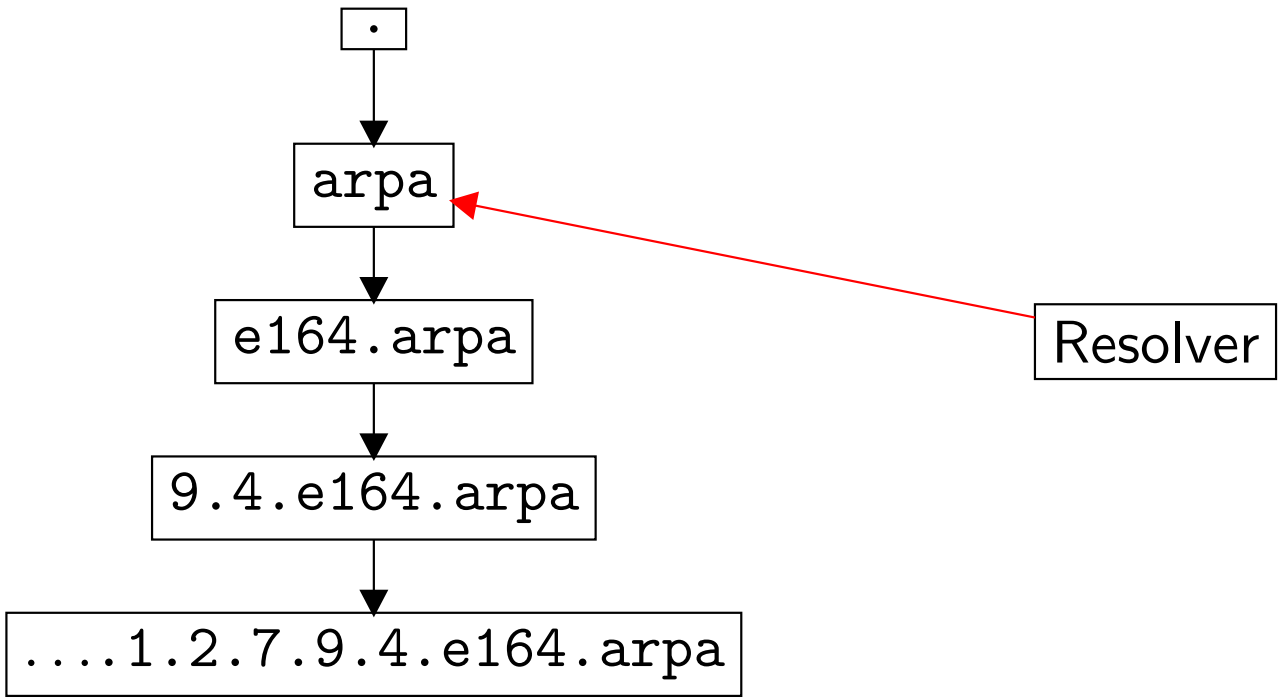
Data leaks along the DNS delegation tree

- Cold cache behavior



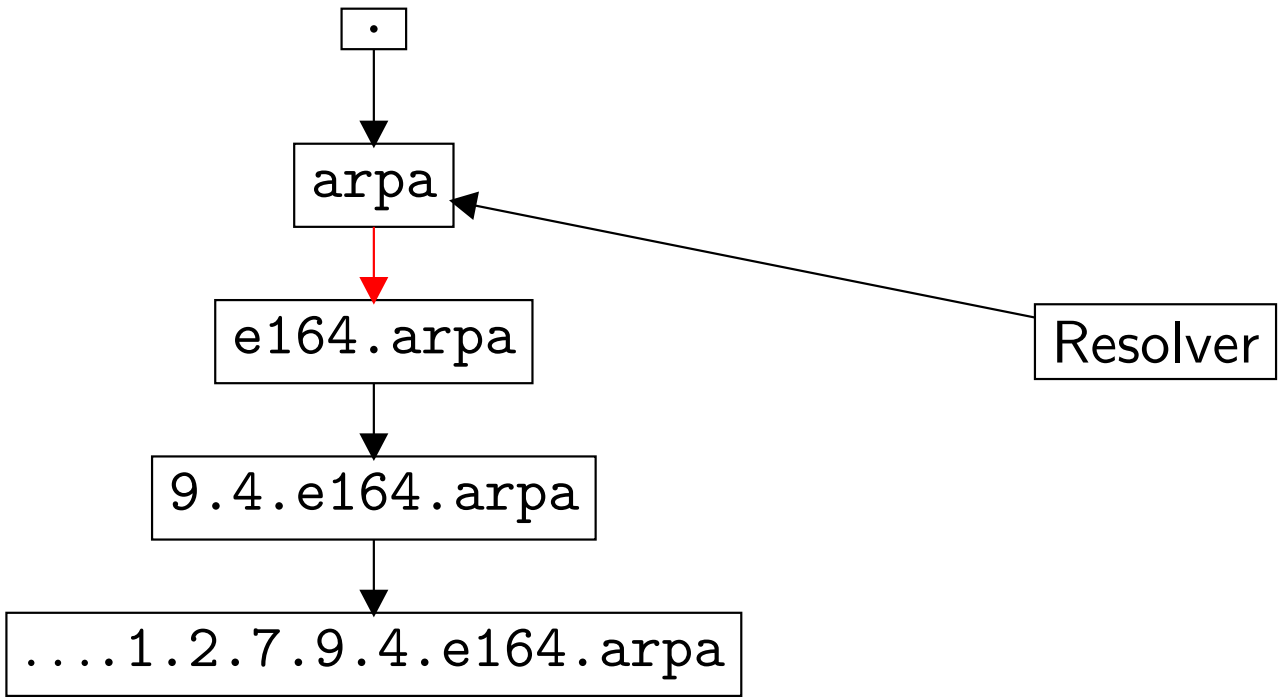
Data leaks along the DNS delegation tree

- Cold cache behavior



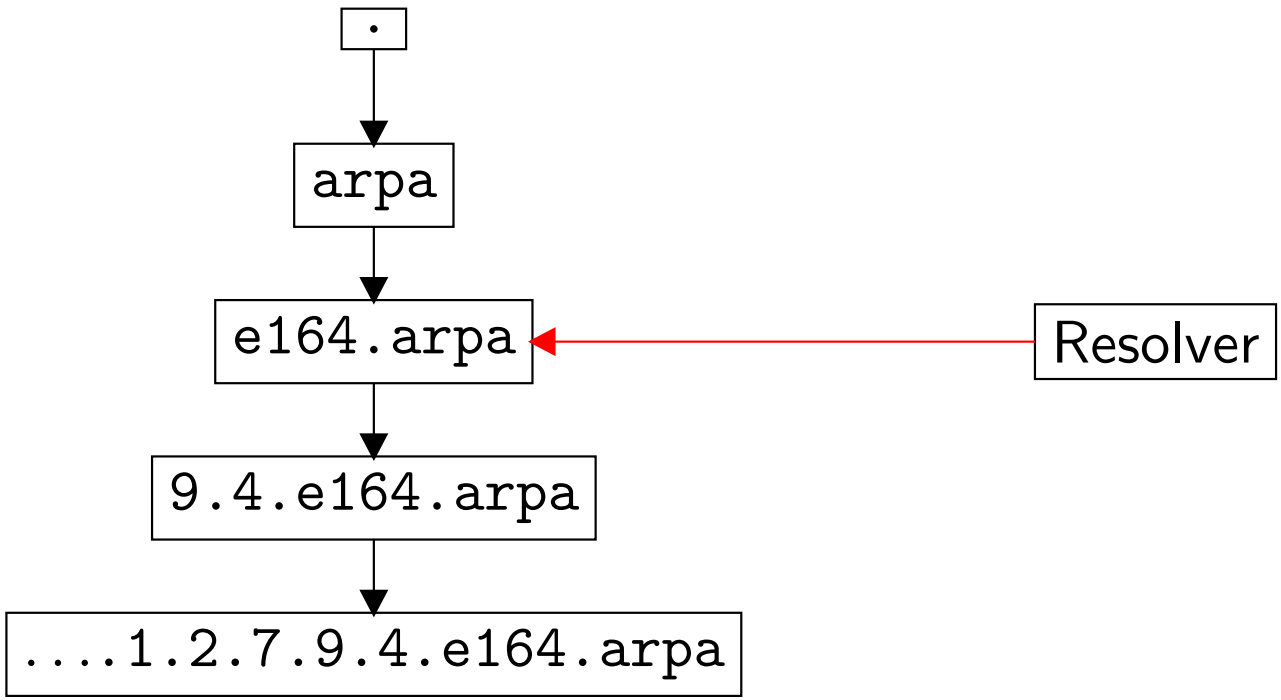
Data leaks along the DNS delegation tree

- Cold cache behavior



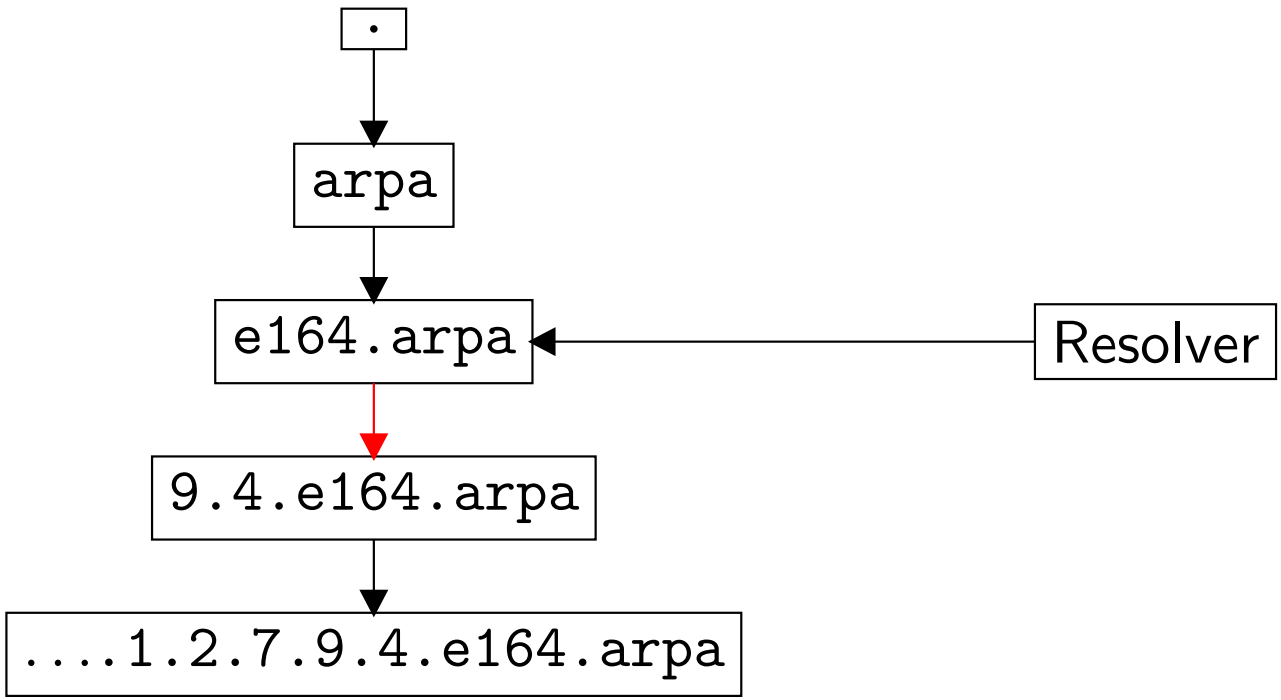
Data leaks along the DNS delegation tree

- Cold cache behavior



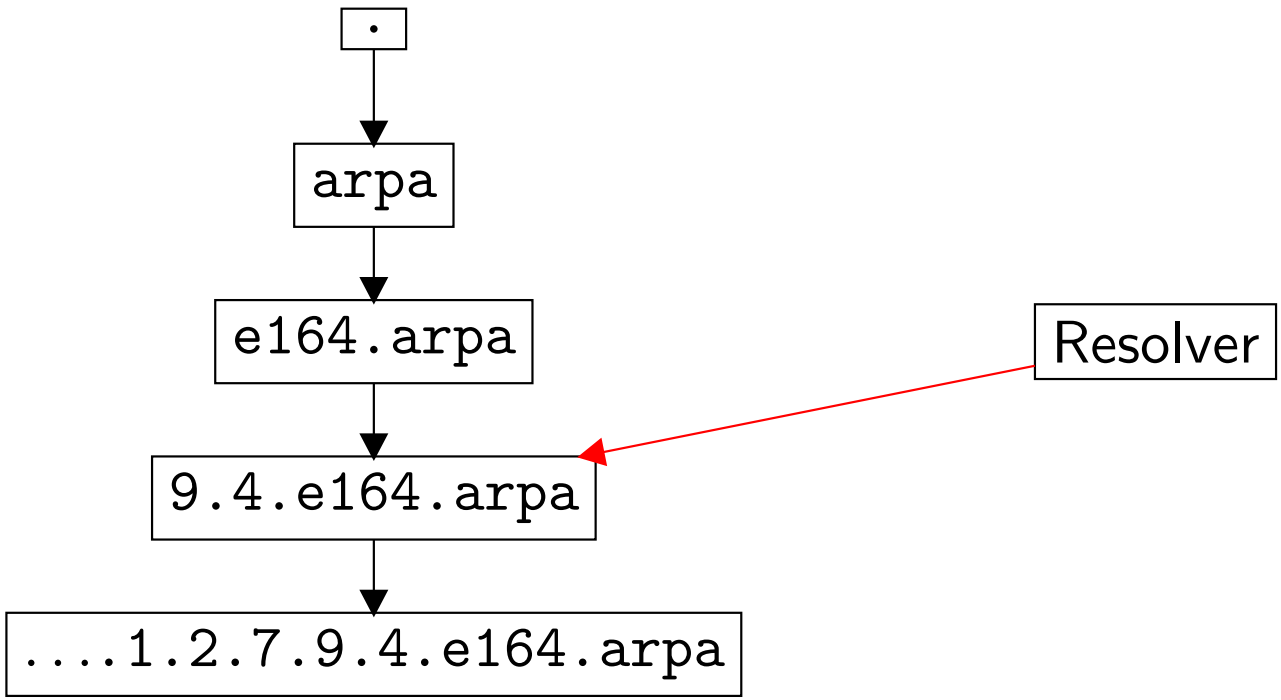
Data leaks along the DNS delegation tree

- Cold cache behavior



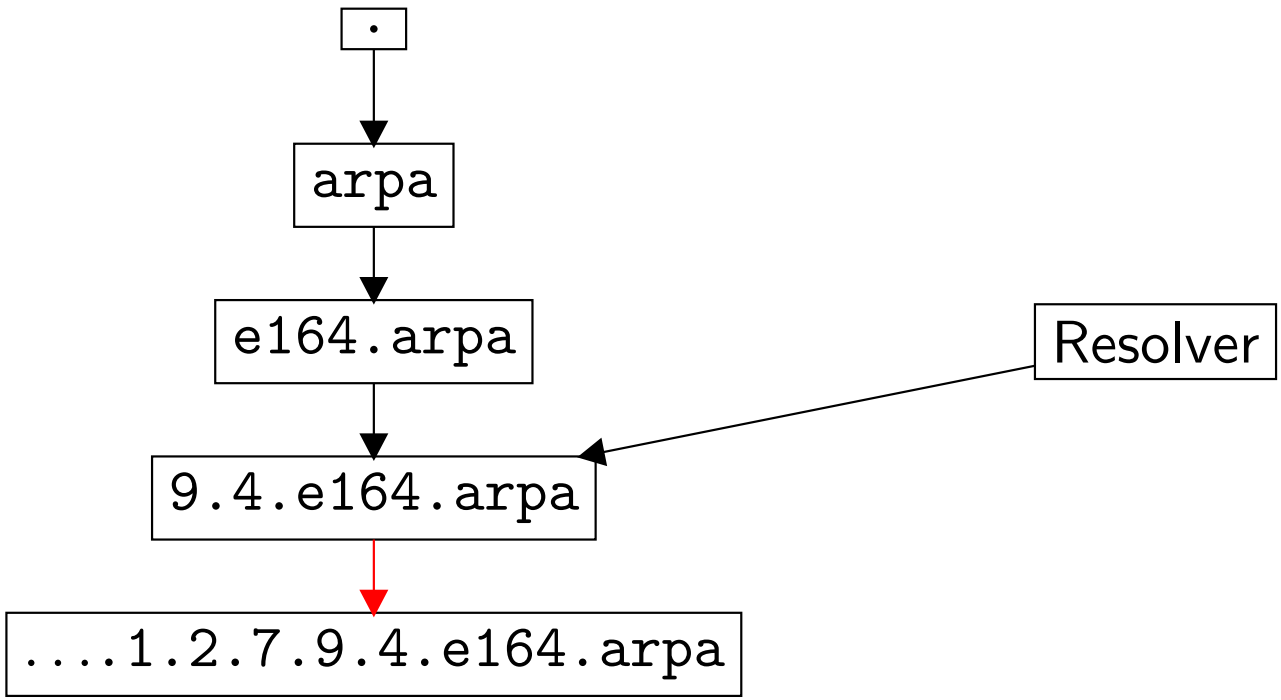
Data leaks along the DNS delegation tree

- Cold cache behavior



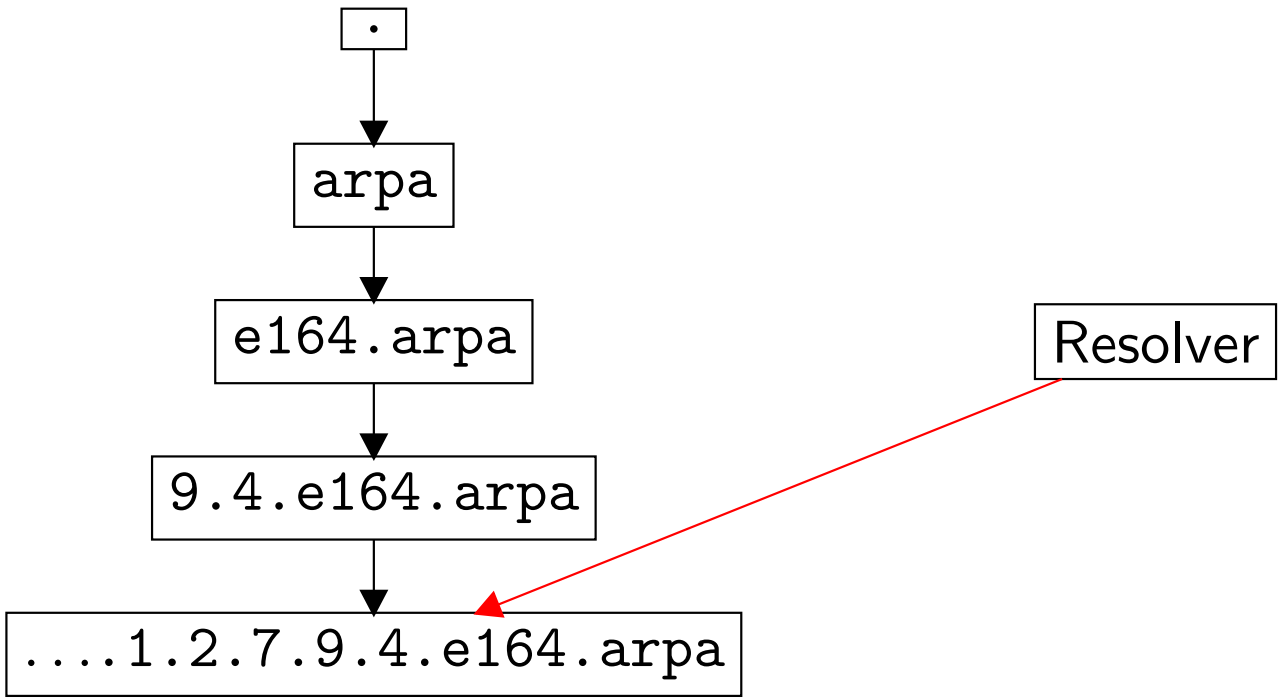
Data leaks along the DNS delegation tree

- Cold cache behavior



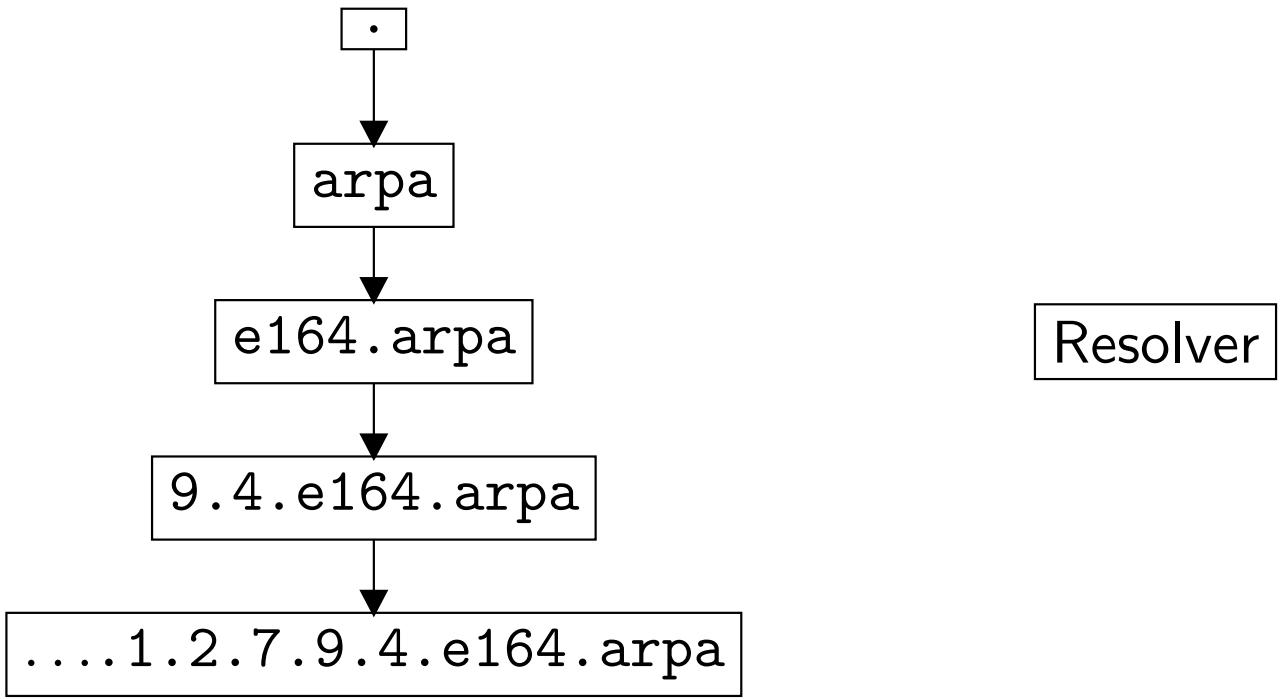
Data leaks along the DNS delegation tree

- Cold cache behavior



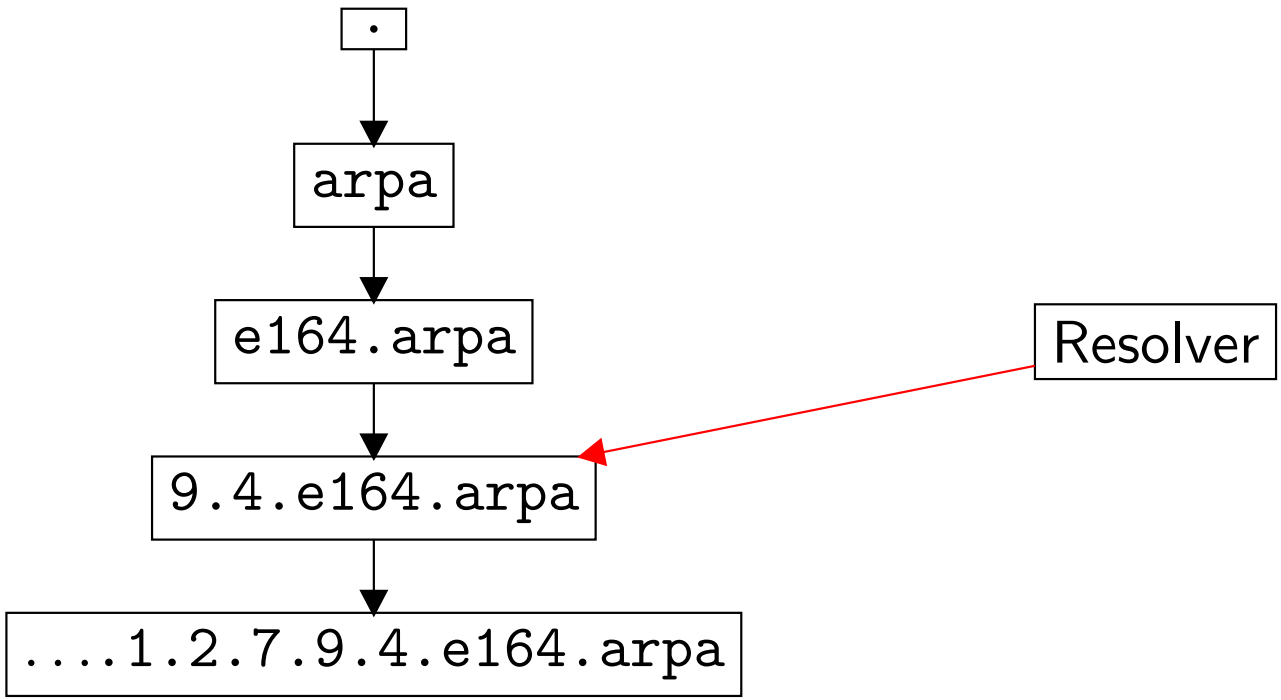
Data leaks along the DNS delegation tree (2)

- Warm cache behavior (country code delegation cached)



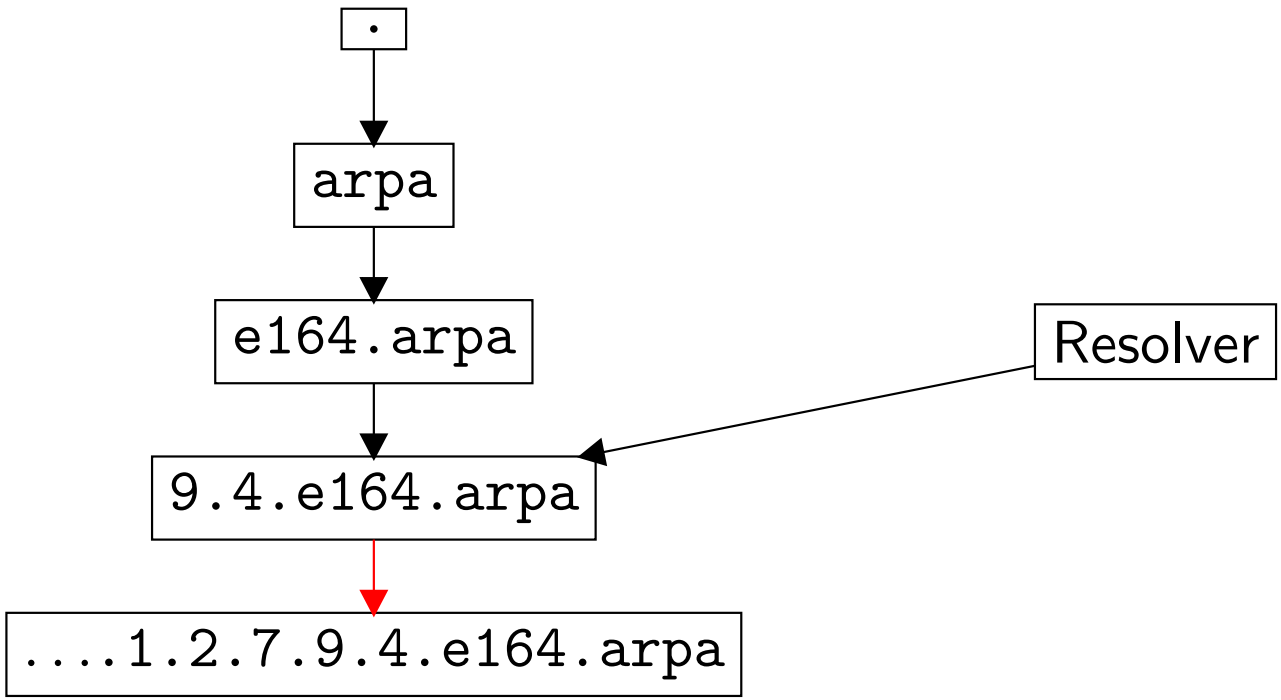
Data leaks along the DNS delegation tree (2)

- Warm cache behavior (country code delegation cached)



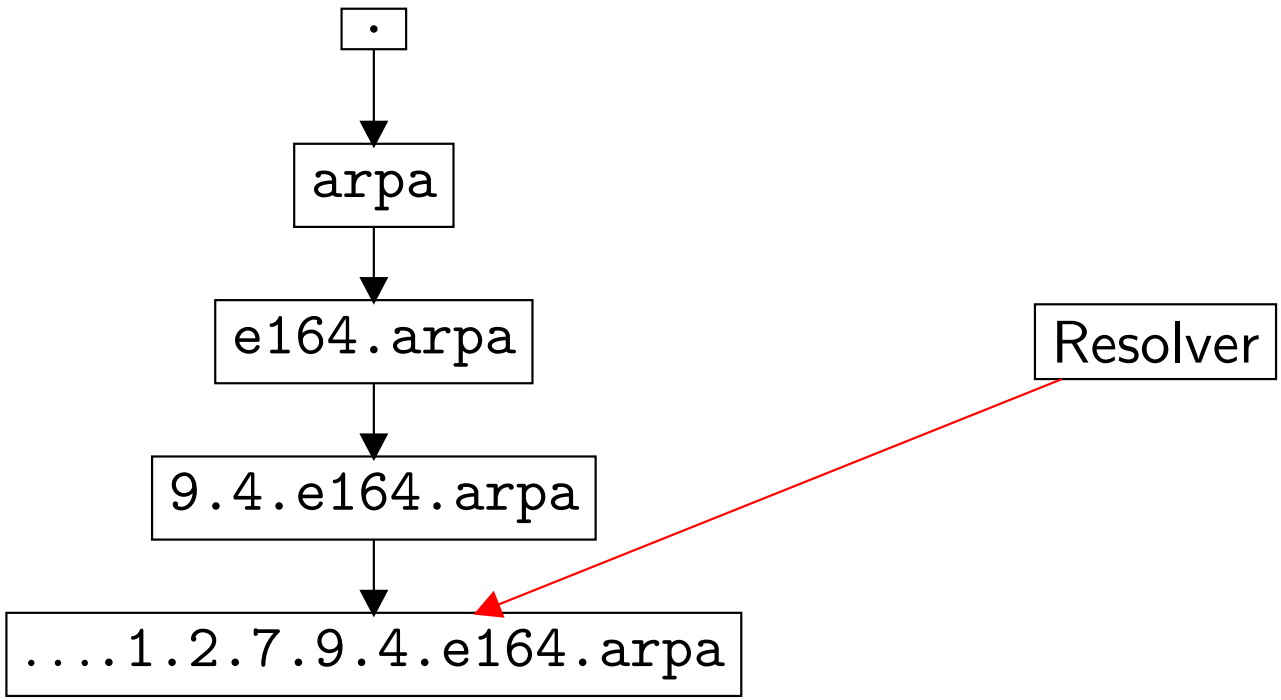
Data leaks along the DNS delegation tree (2)

- Warm cache behavior (country code delegation cached)



Data leaks along the DNS delegation tree (2)

- Warm cache behavior (country code delegation cached)



Data leaks along the DNS delegation tree (3)

- Each authoritative server receives essentially the same query.
- *The complete called number is included in each query.*
- Once the country code delegation is cached, the data longer leaks to the `e164.arpa` servers.
- The `e164.arpa` servers are located in US, JP, SE, NL, CN – are you sure no one is mining your calls?
- Maybe you should get that country code delegation after all.

Outline

Confidentiality of the subscriber list

- Phone numbers are personally identifiable information. Unlisted and unpublished phone numbers exist.
- Some `e164.arpa` registries operate WHOIS services which contain additional subscriber information (names, addresses).
- Registrars might share a business interest in keeping their market coverage secret.

Brute force zone enumeration

- The brute force approach needs about 10^{10} queries for the NANP.
- Factoring in knowledge of the numbering plan can reduce the necessary work by a significant factor.
- This is especially true for countries with variable area code length (Germany, for instance).
- The query volume is substantial. People would notice.
- But we can do much better.

BFK

DNS excursion: empty non-terminals

- *Empty non-terminals* are domain names which exist (because they have subdomains), but have no resource records associated with them.
- Example:
  - `9.4.e164.arpa` exists (it has got NS and SOA records, at the very least),
  - so `4.e164.arpa` exists as well, but there are no resource records for it.
- This definition is arbitrary. See RFC 4592 for some of the implications.

DNS excursion: queries for empty non-terminals

- Queries for empty non-terminals just return an empty record set, not an error.
- On the other hand, a *Name Error* means that the domain does not exist, *including all of its potential subdomains*.
- This means it is possible to detect the presence and absence of entire subtrees in the DNS.

The enumeration algorithm

- Start with some country code delegation:
  - `9.4.e164.arpa`

The enumeration algorithm

- No interesting data, so look at the subdomains.
  - `0.9.4.e164.arpa`
  - `1.9.4.e164.arpa`
  - `2.9.4.e164.arpa`
  - `3.9.4.e164.arpa`
  - `4.9.4.e164.arpa`
  - `5.9.4.e164.arpa`
  - `6.9.4.e164.arpa`
  - `7.9.4.e164.arpa`
  - `8.9.4.e164.arpa`
  - `9.9.4.e164.arpa`

The enumeration algorithm

- No interesting data, so look at the subdomains.
  - `0.9.4.e164.arpa` – does not exist
  - `1.9.4.e164.arpa`
  - `2.9.4.e164.arpa`
  - `3.9.4.e164.arpa`
  - `4.9.4.e164.arpa`
  - `5.9.4.e164.arpa`
  - `6.9.4.e164.arpa`
  - `7.9.4.e164.arpa`
  - `8.9.4.e164.arpa`
  - `9.9.4.e164.arpa`

The enumeration algorithm

- Continue with one of the remaining domains, e. g. `1.9.4.e164.arpa`:
  - `0.1.9.4.e164.arpa`
  - `1.1.9.4.e164.arpa`
  - `2.1.9.4.e164.arpa`
  - `3.1.9.4.e164.arpa`
  - `4.1.9.4.e164.arpa`
  - `5.1.9.4.e164.arpa`
  - `6.1.9.4.e164.arpa`
  - `7.1.9.4.e164.arpa`
  - `8.1.9.4.e164.arpa`
  - `9.1.9.4.e164.arpa`

The enumeration algorithm

- Continue with one of the remaining domains, e. g. `1.9.4.e164.arpa`:
  - `0.1.9.4.e164.arpa` – does not exist
  - `1.1.9.4.e164.arpa` – does not exist
  - `2.1.9.4.e164.arpa` – does not exist
  - `3.1.9.4.e164.arpa` – does not exist
  - `4.1.9.4.e164.arpa` – does not exist
  - `5.1.9.4.e164.arpa`
  - `6.1.9.4.e164.arpa`
  - `7.1.9.4.e164.arpa`
  - `8.1.9.4.e164.arpa`
  - `9.1.9.4.e164.arpa` – does not exist

The enumeration algorithm

- Continue with one of these:
  - `5.1.9.4.e164.arpa`
  - `6.1.9.4.e164.arpa`
  - `7.1.9.4.e164.arpa`
  - `8.1.9.4.e164.arpa`
- And so on.

Enumeration experiments

- Experimental data:

CC	Domains	Queries	Seconds
+43	5,802	97,351	344
+49	6,876	279,281	1,005
+971	1	21	2

  (Timing information is for a crude, multi-threaded Perl implementation running on a cheap, consumer-grade ADSL line.)
- Highly parallelizable.
- Efficiency (queries per actually used domain) increases with the density of the registrations.

Countermeasures

- Never return *Name Error* (but preserve the wildcard semantics).
- NSEC3 (DNSSEC without trivial zone enumeration) seems to expose subtree existence even without explicit RCODEs.
- Synthetic records can likely be detected easily.

Outline

DNS Cache Snooping

- Basic idea:
  - Query recursive resolvers without requesting recursion.
  - Resolvers return whatever data they have in their caches.
- Together with the original TTL, the time of the first request (that is, the phone call) can be established.
- Cache snooping works well for regular domains, too.

Cache Snooping and ENUM

- Only the called party is revealed.
- However, the location of the resolver might provide a hint to the caller.
- The enumeration algorithm described above does not work.
  - During regular operations, the required Name Error responses do not enter the cache.
- Thanks to negative caching, it does not matter if the called party has registered a ENUM domain.

Cache Snooping and ENUM

- Only the called party is revealed.
- However, the location of the resolver might provide a hint to the caller.
- The enumeration algorithm described above does not work.
  - During regular operations, the required Name Error responses do not enter the cache.
- Thanks to negative caching, it does not matter if the called party has registered a ENUM domain.

Countermeasures

- Make your resolvers authoritative for the ENUM domains (the zones are public anyway, see above).
- Use a dedicated full-service resolver for each ENUM-enabled device.
- Anycasting, load-balancing, and not providing DNS resolver services to the whole word may make the attacker's job harder.
- With filtered non-recursive queries, information still leaks through timing differences.

BFK

Summary

- It is possible to enumerate ENUM zones with reasonable effort.
- Caching DNS resolvers may leak call-related information once the caller uses ENUM.
- Nevertheless, it is desirable to obtain at least a country-code delegation to contain some of the data leaks.

Thank you for your attention!

Questions?