DNSSEC and the network issues

Stéphane Bortzmeyer
AFNIC
bortzmeyer@nic.fr

OARC workshop (Denver, USA) - October 2010

Background

Background

1. AFNIC manages six TLDs, for various French outposts. Most are frozen and not actually used. Great for testing! All TLDs have the same set of servers.

Background

1. AFNIC manages six TLDs, for various French outposts. Most are frozen and not actually used. Great for testing! All TLDs have the same set of servers.
2. They were signed from April to September ".FR" was signed last, on September 14th.

Background

1. AFNIC manages six TLDs, for various French outposts. Most are frozen and not actually used. Great for testing! All TLDs have the same set of servers.
2. They were signed from April to September ".FR" was signed last, on September 14th.
3. The signing of small TLDs allowed to discover bugs (BIND bug [ISC-Bugs #22007] with NSEC3 when only one name is signed) and network issues.

Network

But network issues were supposed to be fixed at least since the signing of the root in July!

Unfortunately, no. Two of the authoritative name servers of the set were not able to send back DNSSEC responses in some cases.

Warning

Obviously, the idea is not to shame someone in particular. A name server admin today has **many** things to watch out. Anyone can make mistakes. And DNSSEC is still new.

But I want to emphasize that not everything is settled yet. We still have work to do.

Detection

“As long as you don’t test anything, you get peaceful sleep.”

Unfortunately, some people wrote DNS testing software: http://www.bortzmeyer.org/tests-dns.html

```
% zonecheck pm
...
       -------------
     ,-------------.|
~~~~ |    fatal    || ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     `-------------'
f> Server doesn't listen/answer on port 53 for TCP protocol
 | Ref: IETF RFC1035 (p.32 4.2. Transport)
 |   The DNS assumes that messages will be transmitted as datagrams or
 | a byte stream carried by a virtual circuit. While virtual circuits c
 | be used for any DNS activity, datagrams are preferred for queries du
 | to their lower overhead and better performance.
 `----- -- -- - -  -
=> g.ext.nic.fr/2001:500:14:6039:AD::1
```

Investigation

The above server does work with TCP. But Zonecheck is aggressive: it queries with the DO bit on **and** it asks for ANY data the server has.

SOA answers, even with the DO bit set, were less than 1500 bytes (Ethernet MTU). ANY answers were not.

```
% dig +dnssec +tcp @2001:500:14:6039:ad::1 ANY pm.

; <<>> DiG 9.6-ESV-R1 <<>> +tcp @2001:500:14:6039:ad::1 ANY pm.
; (1 server found)
;; global options: +cmd
;; connection timed out; no servers could be reached
```

Interrogation

Do such queries really happen in the wild? Isn't it an artificial test?

They do not happen often but the idea of a DNS testing tool is to push the server to the limit, to detect problems in advance.

Explanation

For the `2001:500:14:6039:ad::1`, there was a MTU problem in the path. Lowering the MTU on the authoritative server solved the problem.

What is very strange is that the issue, unlike most IPv6 problems, happened with native connections. (Some reported seeing it also with tunnels.) Not fully understood yet.

It is an anycasted server and not all places were affected (mostly in Europe, but also through PAIX in the USA). Complex networking issues.

Repetition

The first problem was easy to detect since it showed itself from our local network.

The second was a bit more difficult to test because it appeared on external networks.

```
%  dig +dnssec -6 @e.ext.nic.fr ANY fr

; <<>> DiG 9.7.1 <<>> +dnssec -6 @e.ext.nic.fr ANY fr
; (1 server found)
;; global options: +cmd
;; connection timed out; no servers could be reached
```

Solution

In this second case, the problem was simpler to understand: the authoritative name server was an OpenBSD machine protected by pf.

There was no rule to allow incoming “too big” ICMP messages.

```
% dig -6 +dnssec @e.ext.nic.fr ANY fr.
...
;; Query time: 109 msec
;; SERVER: 2a00:d78:0:102:193:176:144:6#53(2a00:d78:0:102:193:176:144:6
;; WHEN: Fri Sep 17 11:29:24 2010
;; MSG SIZE  rcvd: 2228
```

Conclusion

Do not spend time making (or listening to) speeches at OARC, check your setup, debug, and check again.

Thanks to PCH and SIDN for their reactivity and their quick solutions.