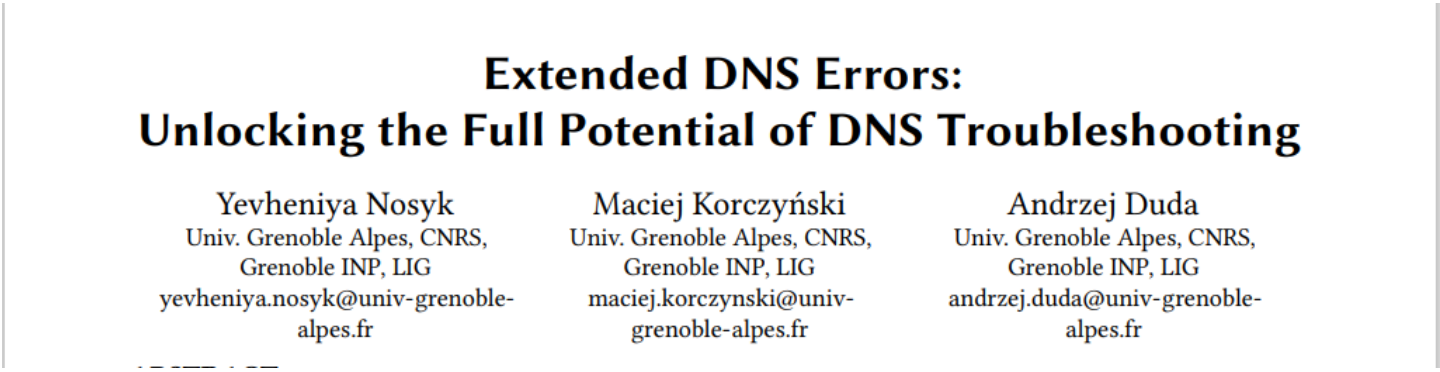
Extended DNS Errors: 释放 DNS 故障排除的全部潜力

引言

本次分享的论文是关于 **Extended DNS Errors（EDE）在定位域名解析故障根因方面的实践效果**。域名解析故障根因的定位始终是一个难题，RFC 8914 (Extended DNS Errors or EDE) 通过在 OPT 资源记录中定义一些新的标志位来解决该问题。作者对支持 EDE 的四个主要 DNS 提供商和三个大型公共 DNS 解析器进行测试，以探究 EDE的实践效果。作者发现，EDE 能够使得管理者缩小 DNS 解析故障根因范围，但针对测试的 DNS 解析器，94% 返回的标志位都不相同，**Cloudflare DNS 是 EDE 标准的最佳实现**（回复的最精确最精确）。因为，作者在 Cloudflare DNS 解析器上进行 303 M 个注册域名的测试，有 17.7 M 个域名触发了 EDE 协议，其中，**无效授权（Lame delegations）和 DNSSEC 认证失败是最常见的 DNS 解析错误**。



Extended DNS Errors:
Unlocking the Full Potential of DNS Troubleshooting

Yevheniya Nosyk
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
yevheniya.nosyk@univ-grenoble-alpes.fr

Maciej Korczyński
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
maciej.korczynski@univ-grenoble-alpes.fr

Andrzej Duda
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
andrzej.duda@univ-grenoble-alpes.fr

背景介绍

Domain Name System (DNS)

域名系统的出现可追溯到 1987 年【1】，用于将人类可读的域名转换为计算机可读的 IP 地址。在 DNS 设计初期，**DNS 数据包头部留有 4 bit 长度的 RCODE 用于回复 DNS 解析错误的情形**，并且直接定义了 6 种情形，剩余的标志位留给后续的扩展。

但随着互联网的发展，DNS 协议变得越来越复杂，**截止到 2023 年，已经有 297 个 RFC 用于定义和规范 DNS 的设计**，但同时也使得 DNS 的误配置和 解析失败的情形越来越多，据测量研究， 仅有 68.1% 的 DNS 解析是成功的【2】。

DNS 故障定位

随着 DNS 协议变得越来越复杂，最初设计的 DNS 协议格式已经远远无法支持新的协议特性。为了解决该问题，**EDNS0（EDNS-Client-Subnet Extension）标准引入新的 OPT 资源记录来提供新的扩展选项（OPTION）**。OPTION-CODE 15 被分配给了 EDE，RFC 8914 提出的标准另外定义了16 bit 的 INFO-CODE 字段用于新添加的错误标志位和可变长度的 EXTRA-TEXT 字段，以更准确的返回域名解析过程中出现的错误。另外，**RFC 8919 中的定义和 RCODE 不冲突，二者可以任意混合使用。**新定义的错误标志位如表 1 所示，前 25 个错误标志（0~24）在 RFC 8919 中直接定义，后 5 个（25~29）标志位随后被添加到 IANA 注册表中。

Code	Description	Code	Description
0.	Other	15.	Blocked
1.	Unsupported DNSKEY Algorithm	16.	Censored
2.	Unsupported DS Digest Type	17.	Filtered
3.	Stale Answer	18.	Prohibited
4.	Forged Answer	19.	Stale NXDOMAIN Answer
5.	DNSSEC Indeterminate	20.	Not Authoritative
6.	DNSSEC Bogus	21.	Not Supported
7.	Signature Expired	22.	No Reachable Authority
8.	Signature Not Yet Valid	23.	Network Error
9.	DNSKEY Missing	24.	Invalid Data
10.	RRSIGs Missing	25.	Signature Expired before Valid
11.	No Zone Key Bit Set	26.	Too Early
12.	NSEC Missing	27.	Unsupported NSEC3 Iter. Value
13.	Cached Error	28.	Unable to conform to policy
14.	Not Ready	29.	Synthesized

表1：已经注册的扩展 DNS 错误标志位

已定义的 INFO-CODE 涉及 DNS 的方方面面，并且适用于转发器，递归解析器，权威服务器等整个 DNS 系统，具体如下：

1. **DNSSEC 验证：**1，2，5-12，25，27
2. **缓存：**3，13，19，29
3. **DNS 解析流程：**4，15-18，20
4. **DNS 软件操作：**14，21，22，23
5. **其他：**0，24，26，28

截至到 2023 年 5 月， DNS 的四大提供商（BIND9，Unbond，PowerDNS Recursor 和 Knot Resolver）已经实现了部分 RFC 8916（已定义的错误标志位的子集）的定义。

触发 EDE

针对域名解析过程中可能出现的错误，作者配置了 63 个故障子域名，具体如表 2 所示，并在 3 个公共 DNS 解析器（Cloudflare DNS，Quad9，和 OpenDNS）和 4 个 DNS 软件提供商的软件（BIND 9.19.9，Knot Resolver 5.6.0，Unbound 1.16.2 和 PowerDNS Recursor 4.8.2.）进行测试以充分的触发 EDE。

故障域名列表为：https://extended-dns-errors.com.

#	Description	Subdomains
1.	Control subdomain	valid
2.	DS misconfigurations	no-ds, ds-bad-tag, ds-bad-key-algo, ds-unassigned-key-algo, ds-reserved-key-algo, ds-unassigned-digest-algo, ds-bogus-digest-value
3.	RRSIG misconfigurations	rrsig-exp-all, rrsig-exp-a, rrsig-not-yet-all, rrsig-not-yet-a, rrsig-no-all, rrsig-exp-before-all, rrsig-no-a, rrsig-exp-before-a
4.	NSEC3 misconfigurations	nsec3-missing, bad-nsec3-hash, bad-nsec3-next, bad-nsec3-rrsig, nsec3-rrsig-missing, nsec3-iter-200, nsec3param-missing, bad-nsec3param-salt, no-nsec3param-nsec3
5.	DNSKEY misconfigurations	no-zsk, bad-zsk, no-ksk, no-rrsig-ksk, bad-rrsig-ksk, bad-ksk, no-rrsig-dnskey, bad-rrsig-dnskey, no-dnskey-256, no-dnskey-257, no-dnskey-256-257, bad-zsk-algo, unassigned-zsk-algo, reserved-zsk-algo
6.	Invalid AAAA glue records	v6-mapped, v6-multicast, v6-unspecified, v4-hex, v6-unique-local, v6-doc, v6-link-local, v6-localhost, v6-mapped-dep, v6-nat64
7.	Invalid A glues records	v4-private-10, v4-doc, v4-private-172, v4-loopback, v4-private-192, v4-reserved, v4-this-host, v4-link-local
8.	Other	unsigned, ed448, rsamd5, dsa, allow-query-none, allow-query-localhost

表 2：自行配置的故障子域名

测试结果

利用预先配置的 63 个故障子域名对 3 个公共 DNS 解析器和 4 个 DNS 软件提供商的软件进行测试，作者发现，只有 4 个子域名（no-ds，nsec3-iter-200 ，unsigned 和 valid）测试返回的结果相同，剩余 59 个故障域名的测试，返回结果均不相同，具体测试信息如表 3 所示。

#	Subdomain	BIND 9.19.9	Unbound 1.16.2	PowerDNS 4.8.2	Knot 5.6.0	Cloudflare DNS	Quad9	OpenDNS
1.	valid	None	None	None	None	None	None	None
2.	no-ds	None	None	None	None	None	None	None
3.	ds-bad-tag	None	9	9	6	9	9	6
4.	ds-bad-key-algo	None	9	9	6	9	9	6
5.	ds-unassigned-key-algo	None	None	None	0	9	None	6
6.	ds-reserved-key-algo	None	None	None	0	1	None	6
7.	ds-unassigned-digest-algo	None	None	None	0	2	None	None
8.	ds-bogus-digest-value	None	9	9	6	6	9	6
9.	rrsig-exp-all	None	7	7	7	7	7	6
10.	rrsig-exp-a	None	6	7	None	7	6	7
11.	rrsig-not-yet-all	None	9	8	8	8	9	6
12.	rrsig-not-yet-a	None	6	8	None	8	8	8
13.	rrsig-no-all	None	10	10	10	10	9	6
14.	rrsig-no-a	None	10	10	10	10	10	None
15.	rrsig-exp-before-all	None	9	7	7	10	9	6
16.	rrsig-exp-before-a	None	6	7	None	7	7	7
17.	nsec3-missing	None	12	None	12	6	None	12
18.	bad-nsec3-hash	None	6	None	6	6	6	12
19.	bad-nsec3-next	None	6	None	6	6	6	6
20.	bad-nsec3-rrsig	None	6	None	6	6	None	6
21.	nsec3-rrsig-missing	None	12	None	10	6	9	12
22.	nsec3param-missing	None	10	10	10	10	9	6
23.	bad-nsec3param-salt	None	12	None	12	6	9	12
24.	no-nsec3param-nsec3	None	10	10	10	10	10	6
25.	nsec3-iter-200	None	None	None	None	None	None	None
26.	no-zsk	None	9	6	6	6	9	6
27.	bad-zsk	None	9	6	6	6	6	6
28.	no-ksk	None	9	9	6	9	9	6
29.	no-rrsig-ksk	None	10	9	6	10	9	6
30.	bad-rrsig-ksk	None	9	6	6	6	6	6
31.	bad-ksk	None	9	9	6	9	9	6
32.	no-rrsig-dnskey	None	10	10	10	10	9	6
33.	bad-rrsig-dnskey	None	9	6	6	6	9	6
34.	no-dnskey-256	None	9	6	6	6	9	6
35.	no-dnskey-257	None	9	9	6	9	9	6
36.	no-dnskey-256-257	None	9	10	10	9	10	6
37.	bad-zsk-algo	None	9	6	6	6	6	6
38.	unassigned-zsk-algo	None	9	6	6	6	9	6
39.	reserved-zsk-algo	None	9	6	6	6	6	6
40-49.	v6-mapped, v6-multicast, v6-unspecified, v4-hex, v6-unique-local, v6-doc, v6-link-local, v6-localhost, v6-mapped-dep, v6-nat64	None	None	None	None	22	None	None
50-57.	v4-private-10, v4-doc, v4-private-172, v4-loopback, v4-private-192, v4-reserved, v4-this-host, v4-link-local	None	None	None	None	22	None	None
58.	unsigned	None	None	None	None	None	None	None
59.	ed448	None	None	None	None	1	None	None
60.	rsamd5	None	None	None	0	1	None	None
61.	dsa	None	None	None	0	1	None	None
62.	allow-query-none	None	None	None	None	9,22,23	None	18
63.	allow-query-localhost	None	None	None	None	9,22,23	None	18

表3：63 个故障子域名测试结果

尽管不同解析器软件对于 EDE 的实现上有差异，但不可否认，EDE 机制的实现均有利于缩小域名解析过程中潜在可能故障的空间。另外，在测试的 7 种 DNS 解析器软件中，Cloudflare DNS 对于 EDE 的实现是最准确的。

全网测量

为了测量 EDE 机制在现实中的实际意义，作者先对包含 4.88 亿个域名条目的域名列表（包括 Centralized Zone
Data Service (CZDS), the Tranco list, passive DNS data from SIE Europe, .se, .nu, .ch, .li top-level do-
main (TLD) zone files accessible via AXFR zone transfers, and Google Certificate Transparency logs）进行过滤，排除已经不存在的域名（NXDOMAIN），之后通过 ZDNS 查询其 A 记录进行测试，

具体测试结果如下：

1. **No Reachable Authority，共 13,965,865 个域名**。当没有可达的权威服务器时，通常会返回 Servfail，Cloudflare DNS 通过引用 EDE 机制，返回的错误类型包括 Network Error (23), DNSKEY Missing
(9), and RRSIGs Missing (10) 等。
2. **网络错误，共 11647551 个域名**。网络错误往往会导致递归解析器无法和其他服务器进行通信，通常会返回 Refused 或 Servfail，往往是由权威服务器导致的，引入 EDE 机制，有 14.8M 不同的域名触发了 Network Error (23) 和 No Reachable Authority (22) 的不同组合。
3. **RRSIGs 缺失，共 2746604 个域名**。该错误往往是由于递归解析器无法获得 DNSSEC 验证所导致的，有 2.47 M 个域名在两个 ccTLD 下。但令人惊讶的是，RRSIGs 缺失并没有导致 DNSSEC 验证失败，作者将该问题报告给了 Cloudflare DNS。
4. **DNSKEY 缺失，共 296643 个域名**。该错误是指在父区域找到的 DS 记录与子区域的任何 DNSKEY 都不匹配的情况，但此错误并不意味在子区域中未找到公钥。在 No Reachable Authority (22) 情况下，由于无法和权威服务器建立交互，因此解析程序无法获取 DNSKEY 资源记录。在 RFC 8914 中，DNSKEY missing（9）指的是DNSKEY 没有相对于 DS 记录进行加密验证的情况。
5. **DNSSEC 伪造，共 82465 个域名**。此错误会导致 DNSSEC 验证为伪造状态。解析器无法以加密方式建立从根到请求区域的信任链。超过 8 万个域名产生了 SERVFAIL。此错误还包括 RRSIG 记录无法验证相应的DNSKEY/A RR、DS 哈希与相应的 KSK 不匹配等。
6. **Invalid Data，共 12268 个域名**。此种错误中，未实现 EDNS0 的权威服务器无法使用 RFC 6891[21]中指定的 FORMERR 进行响应，而是仅仅在响应中不包括 OPT 记录。
7. **不支持 DNSSEC 加密算法，共 8751 个域名**。由于未知加密算法的签名无法验证，故递归解析器会默认忽略未知加密算法的 DNSKEY 。另外，在测试中，作者还发现了不支持密钥对应大小的情况。对于不支持 DNSSEC 加密算法的情况，主要有以下两种情形：
    a. 使用禁止的 DNSSEC 算法（DSA-NSEC3-SHA1 或 DSA 等）
    b. 父区域的 DS 记录对应于子区域中 DNSKEY 记录的密钥标签，但算法编号不匹配
8. **Signature Expired，共 2877 个域名**。域名解析器会检查 RRSIG RR 的 Signature Expiration 字段，以判断是否仍然可以用于构建信任链。有 377 个域名因为包含 Signature Expired，而得到 Servfail 的返回结果。
9. **NSNE Missing，共 1980 个域名**。该错误是指在响应中没有返回不存在的有效证明。例如，在该类错误中，域名缺少 NSEC/NSEC3 记录，以用来验证父区域是否缺少 DS 记录或子区域是否缺少 A 记录。“failed to verify an insecure referral proof for <domain>” 已被加到该类所有的错误中。
10. **Unsupported DS Digest Type，共 62 个域名**。IANA 允许使用两种强制性和两种可选算法来计算 DS 摘要值【3】。由于 Cloudflare DNS 不支持可选的 GOST R 34.11-94 算法，在测试的过程中，共有 54 个域名收到了此扩展错误标志。对于其余 8 域名，它们对应的 DS 记录包含未分配的摘要算法类型（unassigned digest algorithm type，8）。

11. **Stale Answer，共 32 个域名**。该错误是指域名解析器利用过期的缓存数据进行响应。由于权威名称服务器返回 REFUSED RCODE，有 6 个域名的响应为 No Reachable Authority (22) 和 Network error (23)。有 12 个域名的响应为 Stale Answer (3) 和 No Reachable authority (22) 的组合，原因为权威域名服务器没有对域名解析器进行响应。
12. **Signature Not Yet Valid ，共 29 个域名**。此 29 个域名中的 1 个为研究人员的故意配置，另外 28 个域名有两对 DNSSEC 签名，分别为有效和 2045 年开始有效。
13. **Cached Error，共 8 个域名**。域名解析器在搜索 Cache 后直接返回 Servfail，可能的原因是域名解析器在先前对该域名的解析过程中失效。
14. **Other，共 7 个域名**。返回此类错误时带有 "超过迭代限制” 的 EXTRA-TEXT，并导致所有域名的返回为 SERVERFAIL 。

总结

在本文中，作者首次分析了 EDE 在现实中的实现。作者首先自行配置了 63 种带有不同错误的域名，并测试了四家DNS 软件供应商和三家公共 DNS 解析器对 RFC 8914 的实现。结果显示，不同 DNS 软件之间对于 EDE 的实现存在高度的不一致性。之后，作者使用 Cloudflare DNS 对 303M 个域名进行测试，发现其中 17.7M 个（5.8%）触发 EDE 响应。作者认为，EDE 机制的实现都能够在不同程度上更准确定位域名解析的错误，能够帮助 DNS 运营商、域名所有者和最终客户识别和解决 DNS 解析过程中存在的问题。

参考文献

[1] Paul Mockapetris. 1987. Domain names - concepts and facilities. RFC 1034. (1987).

[2] Paweł Foremski, Oliver Gasser, and Giovane C. M. Moura. 2019. DNS Observatory: The Big Picture of the DNS. In IMC.

[3] IANA. 2023. DNSSEC Delegation Signer (DS) Resource Record (RR) Type Digest Algorithms. (May 2023).

文章链接

https://hal.science/hal-04216545/document