DNS/UDPでのIP fragmentationの今後

DNSのUDP fragment packetをdropしてよくなる？

Kazunori Fujiwara, JPRS
fujiwara@jprs.co.jp
JANOG 53 LT

自己紹介

- 氏名:　　藤原和典
- 勤務先:　株式会社日本レジストリサービス(JPRS)
  技術研究部
- 業務内容:DNS関連の研究・開発
- JANOG歴
  - 2に参加した記憶
  - 11 12 15 25 31 37.5 40 で発表
- IETFでの活動 (2004~)
  - ENUMプロトコル: RFC 5483 6116
  - メールアドレスの国際化 :RFC 5504 5825 6856 6857
  - DNS関連の問題提起など
    - RFC 7719, 8499: DNS用語集
    - RFC 8198: DNSSECを用いた名前解決の性能向上
    - draft-ietf-dnsop-avoid-fragmentation: DNSでIP断片化を避ける提案

DNS/UDPのIP Fragmentationは脆弱

- "Fragmentation Considered Poisonous"
  - IEEE Conference on Communications and Network Security, Oct 2013
  - 偽造した第二フラグメントをCA(認証局)に送ってキャッシュ汚染できた
- “Domain Validation++ For MitM-Resilient PKI”
  - ACM SIGSAC Conference on Computer and Communications Security , 2018
  - 偽造した第二フラグメントを遠隔からCAに送ってフルリゾルバをキャッシュ汚染し、不適切な証明書を発行させることができた
- “DNS第一フラグメント便乗攻撃の追検証と対策の検討”
  - Kenya Ota and T. Suzuki
  - The 81st National Convention of IPSJ, March 15, 2019
  - この攻撃が実環境でできることを検証した

Path MTU discoveryは脆弱

- IPv4ではICMP Unreachable, Fragmentation neededを送ってやると、受け取ったサーバから指定したIPアドレスへのpath MTUの値を任意の小さな値に指定できる場合がある
  - 20+8+20+8 (56)バイトのパケット一つ
  - Linux, FreeBSDで確認
  - よくできたOSはパケットを送るときにpath MTUの大きさに断片化する
  - IPv6の場合は1280以上のものしか受け付けない
- 攻撃者がフラグメントサイズを設定できるので、第二フラグメントの注入が容易になる

外側のIPv4ヘッダ
45 00 00 3a 00 00 00 00
40 01 checksum source_addr
destination_addr (攻撃先)

ICMP Header
03 04 checksum MTU値(552)

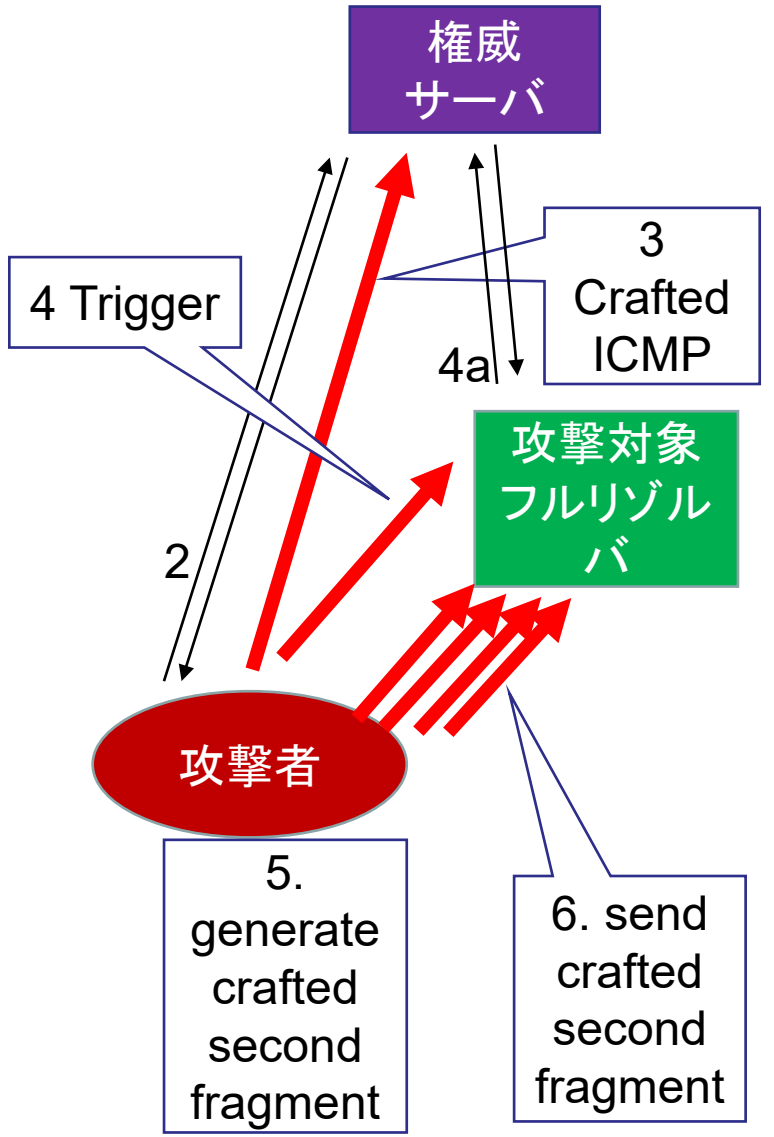
内側のIPv4ヘッダ
45 00 05 78 00 00 40 00
40 11 checksum 攻撃先アドレス
通信先アドレス

内側のUDPヘッダ
00 35 xx xx // source port 53
UDP length 1400
xx xx (checksum field)

論文に書いてある攻撃

1. 攻撃対象を選ぶ (ドメイン名、権威サーバ、フルリゾルバ)
2. 権威サーバからの正しい応答を得る
3. 権威サーバへICMPパケットを送り、リゾルバへのMTUを設定する
4. 攻撃対象フルリゾルバに注入したいドメイン名のクエリを送る

   → 攻撃対象は権威サーバにクエリを送る
5. 攻撃者は第二フラグメントを生成する
6. 攻撃者はIP IDのみ変更した第二フラグメントを権威サーバからの応答が届く前に攻撃対象に65536個送る
   - IP IDは16ビット,65536通りなので成功する可能性がある



IP Fragmentationを使った攻撃への対策

- 解決策は、DNSSEC検証、またはIP Fragmentationを避けること

→ DNSSEC対応は(別途)行うとして、
IP Fragmentationを避ける提案を進めました

DNS/UDPではIP Fragmentationを使うのを避けましょう

- IETF dnsop WGで標準化作業中
  - draft-ietf-dnsop-avoid-fragmentation
    - 著者: Kazunori Fujiwara, Paul Vixie
  - Interface MTUに入りきるようにパケットを作りましょう
    - これでIPv6ではFragmentを使わなくなる
  - IPv4ではDon't Fragment (DF)ビットをセットしてもよい(MAY)
    - SHOULDにしたいけど、いまはまだ無理 (将来的には SHOULD にしたい)
  - 受信側では、Fragmentされた応答を捨ててもよい (MAY)
  - DNS/UDPで扱うデータサイズを1400以下としましょう
    - BIND 9ではmax-udp-size, edns-udp-size
    - (多くのソフトウェアのdefaultは1232なので、すでに1400以下)
    - 超える場合はTruncation bitがセットされる → TCPで再問合せ
- 現在、dnsop WGでの議論は完了し、IESG評価中

この提案がRFCになったら

- 堂々とUDP port 53関連のfragment packetをdropしてもよくなる
  - フルリゾルバの手前で、source port 53/UDPからフルリゾルバへの通信のうち、Fragment Offsetが0, More Fragmentsが1のもの
  - フルリゾルバ専用のIPアドレスなら、第二フラグメント以降を捨ててよい
    - サーバのReassembly bufferにたまるパケットを減らせる
    - フルリゾルバあてで、Fragment Offsetが0でないもの

(発行されなくても、設定してもよいです)
(TCP非対応の権威サーバがあると名前解決に失敗する可能性
→ 攻撃されて誘導されるよりよい)