フルリゾルバ運用者のためのDoH/DoT

version 11.28

SNS投稿NG

株式会社インターネットイニシアティブ
其田　学

Ongoing Innovation

Internet Initiative Japan

自己紹介

Manabu Sonoda
其田 学

所属

株式会社インターネットイニシアティブ
日本DNSオペレーターズグループ 幹事

経歴

- 2008年 AS4704でL1-L8まで行うフルスタックエンジニア
- 2014年 IIJにて現職
  - IIJのお客様提供用のDNSサーバの設計、構築、運用
  - D.DNS.JPの構築、運用
  - コミュニティ活動、啓蒙活動などなど（イマココ）

アジェンダ

- IIJのDNS暗号化への取り組みの紹介
- クライアントの設定
- IIJ Public DNS運用から得られた知見の紹介
- 実サービスへの適用の課題

IIJのDNS暗号化への取り組みの紹介

IIJ Public DNSサービス（ベータ版）をリリース

IIJ、「DNS over TLS」、「DNS over HTTPS」を利用したDNSの試験サービス「IIJ Public DNSサービス（ベータ版）」を提供開始

2019年5月8日

> このニュースのPDF版 [176KB]

株式会社インターネットイニシアティブ（IIJ、本社：東京都千代田区、代表取締役社長：勝 栄二郎、コード番号：3774 東証第一部）は、DNSサーバとの通信を暗号化する「DNS over TLS（DoT）」および「DNS over HTTPS （DoH）」を利用した「IIJ Public DNSサービス（ベータ版）」を、本日より無償公開いたします。本サービスはDNSキャッシュサーバの機能を試験サービスとして提供するもので、DoT/DoHに対応したブラウザ、端末にDoT ホスト名/DoH URLを設定することで、どなたでもご利用いただけます。なお、試験サービスの提供期間は2022年3月31日までを予定しています。

IIJのDNS暗号化への取り組みの紹介

2018年4月

- 既存の実装の調査

2018年10月

- 試験環境下で、構成検討
- 負荷試験を実施

2018年12月

- 試験サービスとして、
  Public DNSを出すことを検討

IIJのDNS暗号化への取り組みの紹介

2019年3月

- IIJ Public DNS構築開始

2019年5月

- IIJ Public DNS提供開始

2020年？

- 正式サービス開始予定

IIJ Public DNSの構成

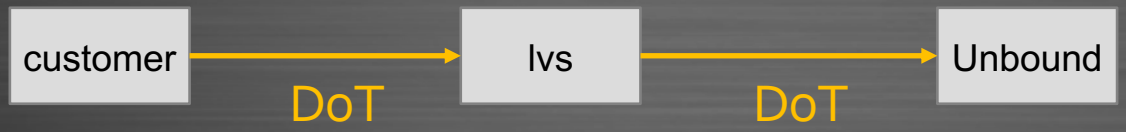
拠点

- 日本
- 米国
- ヨーロッパ

複数箇所で同一IPをAnycast

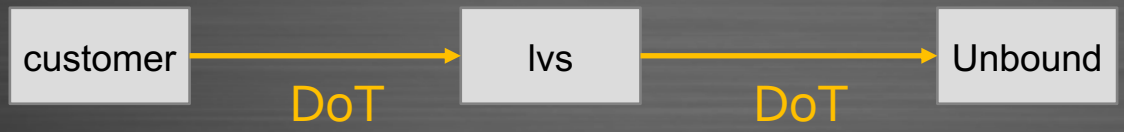
IIJ Public DNSの構成 - DoT

- **UnboundがTLSを終端します。**
- この構成を採用するために、UnboundにDoT関連のパッチを送っています。(1.9でmerge)
  - TLS Session Ticket対応
  - cipher設定対応



IIJ Public DNSの構成 - DoT

- メリット
  - LB側でTLSを解く構成に比べて、TLSを解ける台数を稼げる
  - スケールアウト、スケールアップが用意
    - Unboundホストを増設
    - HWアクセサレータを入れる
  - RateLimitがUnbound側でできる
- デメリット
  - 証明書更新が大変になる
    - （自動化はしますが。。。。。）

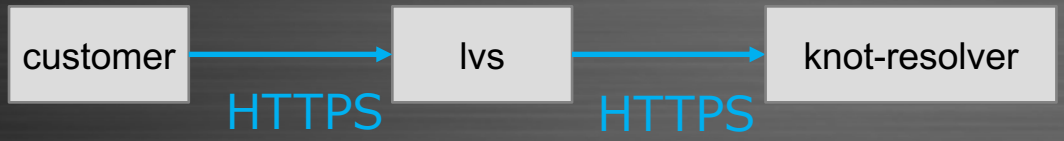


IIJ Public DNSの構成 - DoH

- Unbound はDoH未サポート
- NginxでHTTPS（TLS）終端を行い
- dot-proxyを通してUnboundにクエリーを送る構成

IIJ Public DNSの構成 – DoH （if knot-resolver used)

- knot-resolverはDoH対応しているので、直受けが可能
- 主要なフルリゾルバ実装はDoH対応すると表明しているので、数年後にはこの構成が一般的になると思われる



（おさらい）TraditionalなDNSとの違い

TraditionalなDNSとの違い

	DoUDP	DoTCP	DoT	DoH
トランスポート	UDP	TCP	TCP/TLS	TCP/HTTPS
pipelineing	なし	あり	あり	なし
暗号化			TLS	TLS

- 一番の違いはトランスポートがTCPであること
  - ステートフルになり、ロードバランスはステートを考える必要が出てくる
  - DoTはpipelining、DoHはkeepaliveで接続しっぱなしが基本
- TLSでの暗号化を行う
  - コネクション時の処理が桁違いに重い
  - DSCなど既存のPCAPツールベースのツールが使えなくなる

クライアントの設定

クライアント側の設定

現在（２０１９年）の所、自動設定の標準的な方法はない
基本は手動設定だが、独自に自動設定を実装している

Android (9以降)

OSに設定されたTraditionalなDNS（手動自動問わず）の
IPアドレスに対しDoTが有効か確認し、
有効であれば使用する

Chrome (79以降)

Chrome側で、TraditionalなDNSのIPとDoHのURLの
Mappingを保持しており、それに応じて自動設定される。
自動設定させるには、ソースコードに取り込まれることが必要

Firefox

日本ではデフォルトONになっていないが、
ONにする動きがある。
デフォルトは1.1.1.1(cloudflare)

無効化したい場合は、TraditionalなDNSで
use-application-dns.netのAの応答に空を返せば
無効化できる。
（日本のISPでこれをやったらブロッキングに。。）

Public DNS運用から得られた知見の紹介

Source Hashベースでのロードバランシングが必須

同じコネクションは必ず同じホストに、
終端する必要がある。

ECMPやLBのアルゴリズムとして、
Source アドレスをベースにnext hopを決める必要がある
（Source port & source addressだとなお良い）

TCP Keepalive

DoTは基本つなぎっぱなしだが、idle timeoutの時間は
クライアントによってまちまち

- 一番対応数が多い、Androidの場合

  idle timeout 15s
  Keepalive 3, interval 5

サーバ側でも最低でも３０秒程度のidle timeoutがないと
再接続が大量に発生し、クライアント側の体感速度が落ちる

証明書のアルゴリズム

RSA2048bitはサーバ側の負荷が大きい

- 証明書をEC P256に変えるだけで
  サーバ側のコネクション時のパフォーマンスは２倍上がる
  - (ただし、クライアント側の検証コストは上がる)

- 発行できる証明局
  - Let’s encrypt
  - Globalsign
    - Wildcard証明書有り
  - DigiCert Secure Server (旧 Symantec, Verisign)
    - (サーバ毎ライセンスなので使いにくい)

運用面 ACL Ratelimit

- DoT,DoHをフルリゾルバで終端
  - いままで通りの設定

- DoT、DoHをLBやWEBサーバで終端
  - フルリゾルバには、LBやWEBサーバのIPで問い合わせ
  - LBやWEBサーバソフト側でACLやRatelimitを
    かける必要がる。

ACLに関してはリフレクター攻撃には使われないため、
ACLかけない選択肢もあり

運用面

クエリーログ
- PCAPベースのクエリは使えなくなる
  - DNSTAPや、リゾルバのクエリログ

統計情報
- DNSソフトウェアから取るStat情報は変わらず
- PCAPベースのものは使えなくなる。
  - DSCが使えなくなります！

実サービス適用への課題

レイテンシ

固定回線では問題にならないが、モバイル回線では
致命的な影響

- 接続時の名前解決がタイムアウトしてしまう
- 朝方、昼間の混み合った時間帯

通信に時間がかかる

- DNSの名前解決がタイムアウト
- セッションのタイムアウト

TLS再接続が発生

レイテンシ – 測定

環境

IIJ Mobile+Essential Phone PH-1 (Android 9)

名前解決方法

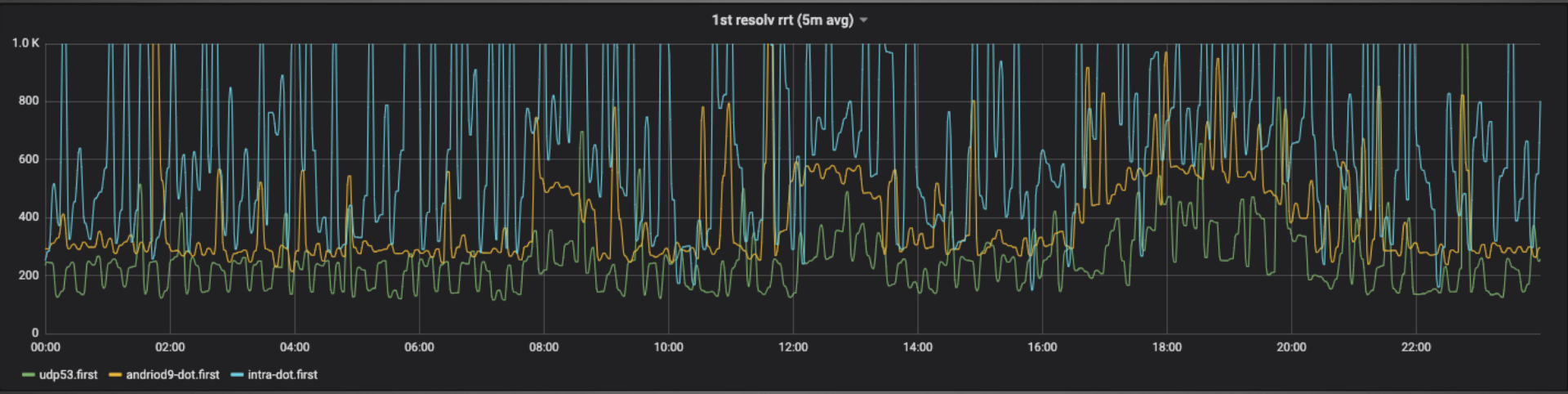
- Traditional DNS
- Android9 Private DNS (DoT)
- Intra(DoH)

測定内容

専用アプリを使って、
Public DNS内のローカルゾーン（測定用ゾーン）に対して
10分毎に、異なるランダムサブドメインのクエリーを２回
投げそのRTTを計測する

TraditionalなDNSとAndroid実装のDoTとのRTTの比較

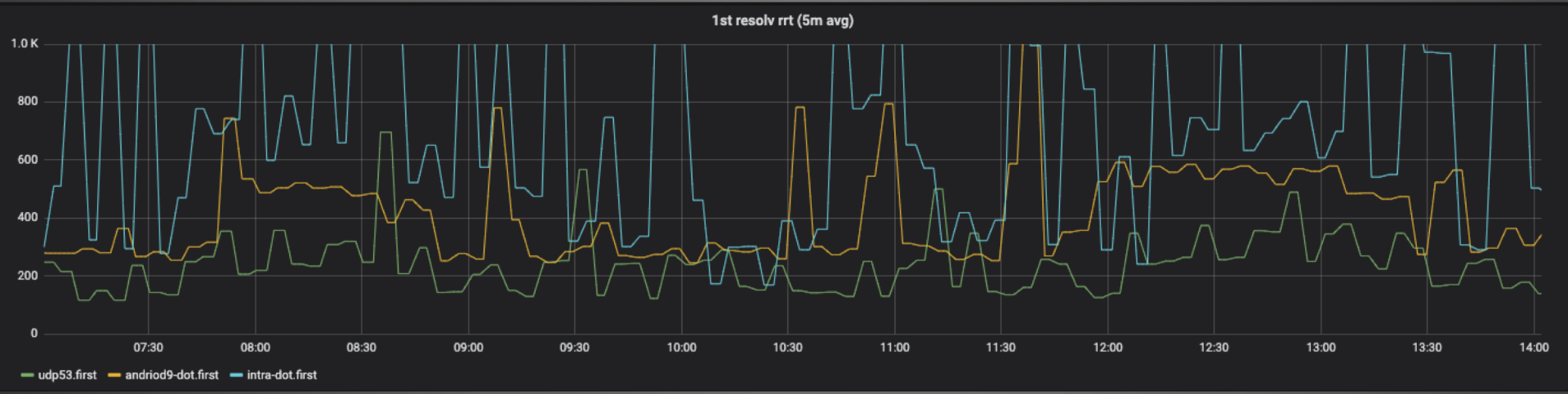
１回目　TLS接続が走るパターン



Android9 UDP
Android9 DoT
Android9+Intra DoH

TraditionalなDNSとAndroid実装のDoTとのRTTの比較

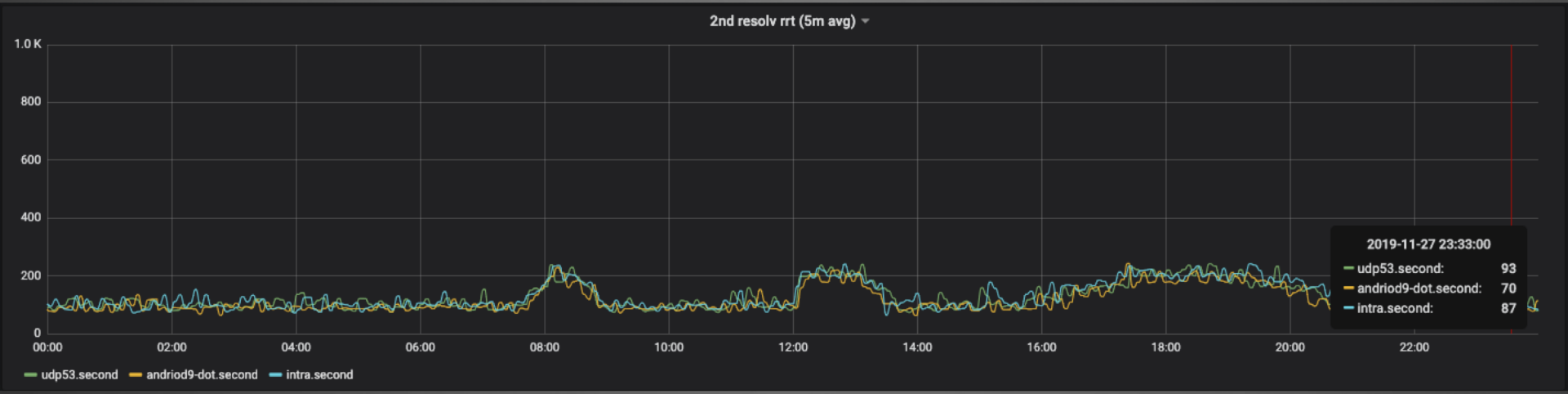
１回目　TLS接続が走るパターン



Android9 UDP
Android9 DoT
Android9+Intra DoH

TraditionalなDNSとAndroid実装のDoTとのRTTの比較

２回目のクエリー



Android9 UDP
Android9 DoT
Android9+Intra DoH

レイテンシ

- UDPですでに200ms越え
- DoTの１回目のクエリーが１秒を超えている場合がある
  - アプリケーションによってはタイムアウトする

- ユーザがブラウザやアプリケーションを立ち上げた時のクエリーがタイムアウトを起こする

- 使い物にならない

レイテンシの改善

今の仕様でできること
- TCP Fast Open
- TLSセッション再開
- TLSv1.3

TCP Fast Openや
セッション再開は

一度接続した後に
再接続を高速にする技術

今回の事象の解決にはならない

クライアント側のTLSv1.3の
普及に期待

レイテンシの改善

本命技術

- HTTP3ベースのDoH
- DNS over QUIC

正直TCPでやるの辛いので、QUICに期待しております

まとめ

- モバイルサービスに適用するには次期尚早
  - レイテンシの改善がないと厳しい
- 固定回線系サービスに適用するには問題ない

Ongoing Innovation

IIJ Internet Initiative Japan

ご清聴ありがとうございました