DNS Summer Day 2025

ECS(RFC 7871)の3社共同検証の結果共有

～商用利用を前提とした検証結果の共有～
※権威DNSサーバー編

株式会社Ｊストリーム　高見澤信弘

2025年6月27日

自己紹介

▶名前：高見澤信弘

▶出身地：山形県天童市

▶所属：株式会社Ｊストリーム (AS24253)

- 新卒でＪストリームへ入社
- エンジニアリング推進室＆プロダクト企画部（アーキテクト）

▶お仕事

- CDN(Content Delivery Network)の企画、構築
- ネットワーク企画

▶好きなもの

- ロードバランサー　→　家にBIG-IP
- おうち19インチラック勢

- 活動
  - IPoE協議会 IPv6地理情報共有推進委員会 幹事
  - 海賊版対策実務者意見交換会 海賊版対策技術検証チーム(WG) メンバー

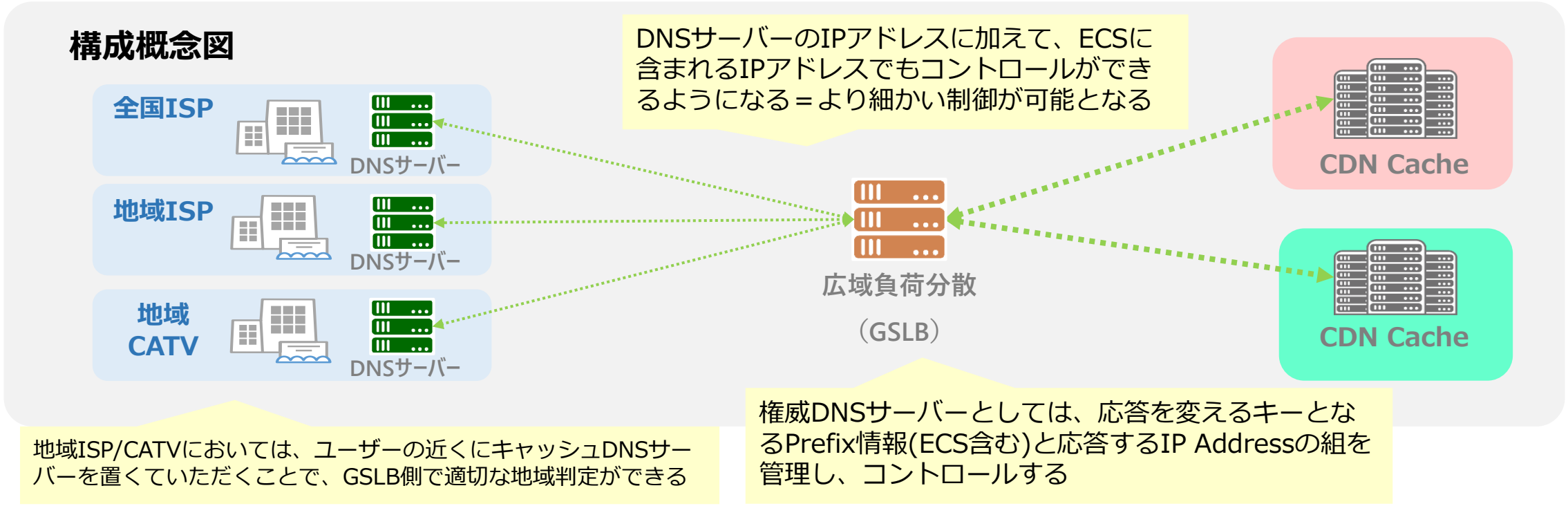
もっと素敵な伝え方を。

Jstream

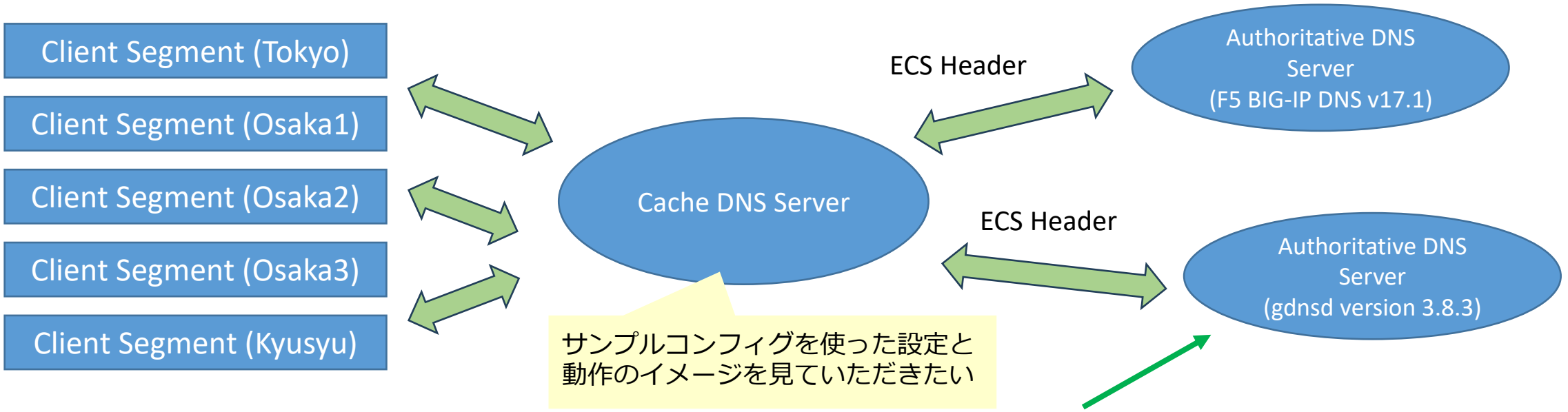
コンテンツ配信側から見たDNSサーバー

▶求める役割：広域負荷分散(Global Server Load Balancing)

- 複数のサーバーへの負荷分散
- サーバーやデータセンターの死活監視
- トラヒックコントロール：キャッシュDNSサーバーのIPを使った応答



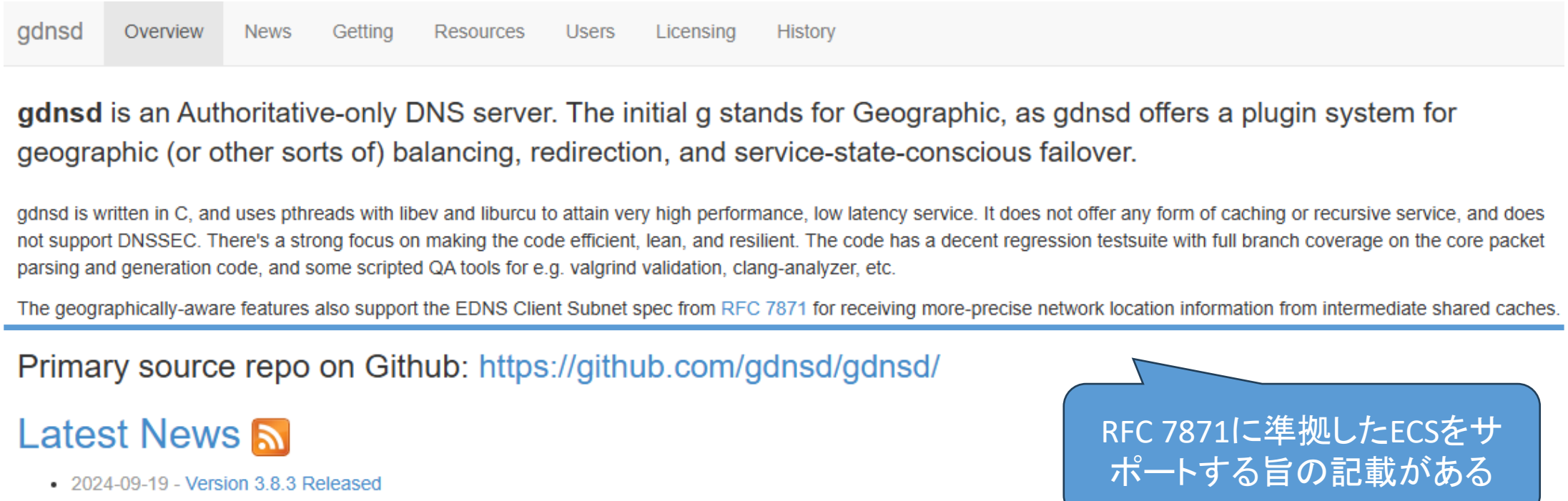
システム構成概要：名前解決動作について



想定地域	ECS(Subnet)	Anser(ServerIP)
大阪1:Osaka1	240b:0010:0200::/41	2001:db8:10::1/128
大阪2:Osaka2	240b:0010:0280::/41	2001:db8:10::2/128
大阪3:Osaka3	240b:0010:0300::/40	2001:db8:10::5/128
九州:Kyusyu	240b:0010:0400::/39	2001:db8:20::1/128
それ以外:Tokyo	::/0 (default)	2001:db8::1/128

オープンソース実装：gdnsd

- gdnsd：https://gdnsd.org/
  - 権威DNSサーバー専用のオープンソースなDNSサーバーの実装
  - ヘルスチェックを使った監視や応答の切り替えなども可能
  - ジオロケーションデータを参照するモジュールもあり、こちらを利用した



RFC 7871に準拠したECSをサポートする旨の記載がある

実際のコンフィグ：gdnsd

- 設定としては
  - Conf：ロードバランスなどの設定
  - Zone：confファイルで設定したリソース名をどのリソースレコードに紐づけるかを指定

```
# cat zones/example.com
$TTL 86400

@    SOA ns1 hostmaster (
     2016030300  ; serial
         7200  ; refresh
         30M   ; retry
         3D    ; expire
         10    ; ncache
)

@    NS    ns1
ns1  A     192.168.1.69

www    10     DYNA    geoip!prod_www
```

右のコンフィグをwwwに紐づける

gdnsd : GdnsdPluginGeoip,
https://github.com/gdnsd/gdnsd/wiki/GdnsdPluginGeoip
(Visited:2025-06-15)

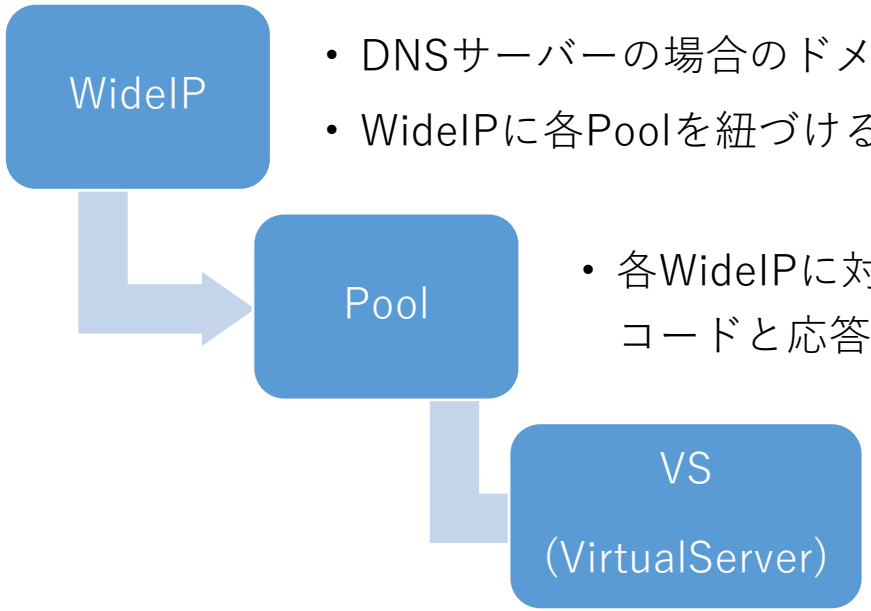
各subnetの宛先を定義

各宛先の回答(IP address)を定義

```
options => {
  edns_client_subnet => true
}
plugins =>{
  geoip => {
    maps => {
      my_prod_map => {
        datacenters => [ dc-01, dc-02, dc-03, dc-04, default ],
         nets => {
           240b:0010:0200::/41 => [ dc-01 ],
           240b:0010:0280::/41 => [ dc-02 ],
           240b:0010:0300::/40 => [ dc-03 ],
           240b:0010:0400::/39 => [ dc-04 ],
           ::/0  => [ default ],
        }
      }
    }
    resources => {
      prod_www => {
        map => my_prod_map
        dcmap => {
          dc-01 => [2001:db8:10::1 ],
          dc-02 => [2001:db8:10::2 ],
          dc-03 => [2001:db8:10::5 ],
          dc-04 => [2001:db8:20::1 ],
          default => [2001:db8::1 ],
      }
    }
  }
}
}
```

アプライアンス：f5 BIG-IP DNS

- GSLBとしては老舗：2000年代初頭には使われていたと思われる
- SLBとGSLBを連携させた監視やロードバランシングが可能
- 設定は以下のように階層化されている



- DNSサーバーの場合のドメイン名設定（FQDNの指定）
- WideIPに各Poolを紐づけることで各リソースレコードを表現する

- 各WideIPに対応するA, AAAA, CNAMEなどの具体的なレコードと応答するIP(Virtual Server)を紐づける

- 監視しているサーバーのリスト

Pool内に複数のサーバーを設定し、その中からECSが指定するPrefix情報をキーとして、どのサーバーを選択させるか設定した。
A,AAAAだけでなく、CNAMEも設定可能

実際のコンフィグ：F5 BIG-IP DNS

Prefix情報と応答するIP Addressを指定

- 全体の設定として、ECSを使ったロードバランスを行うように設定する
- 各subnet事に応答するPool（この場合はIPアドレス）を設定する
  - ロードバランスメソッドとしてtopologyという機能を利用する
  - IPアドレスではなくRegion(Pool)やCNAMEレコードも設定が可能
- これ以外にFQDN設定や監視設定などの設定が必要

```
gtm global-settings load-balancing {
    topology-prefer-edns0-client-subnet enabled
}
gtm topology ldns: subnet 240b:0010:0200::/41 server: subnet 2001:db8:10::1/128 {
    order 1
    score 100
}
gtm topology ldns: subnet 240b:0010:0280::/41 server: subnet 2001:db8:10::2/128 {
    order 2
    score 100
}
Gtm topology ldns: subnet 240b:0010:0300::/40 server: subnet 2001:db8:10::5/128 {
    order 3
    score 100
}
gtm topology ldns: subnet 240b:0010:0400::/39 server: subnet 2001:db8:20::1/128 {
    order 4
    score 100
}
gtm topology ldns: subnet ::/0 server: subnet 2001:db8::1/128 {
    order 13
    score 10
}
```

f5 DevCentral Community, Using Client Subnet in DNS Requests, https://community.f5.com/kb/technicalarticles/using-client-subnet-in-dns-requests/282196 (Visited:2025-06-15)

名前解決動作についての詳細

No.	ECS有無	キャッシュDNSサーバーがリクエストするPrefix	権威サーバー設定	応答内容(AAAA / Scope Subnet)	
1	なし	240b:0010:0000::/40/0	Prefix : ::/0 (default) AAAA : 2001:db8::1	2001:db8::1	240b:0010:0000::/40/39
2	あり	240b:0010:0200::/40/0	Prefix : 240b:0010:0200::/41 AAAA : 2001:db8:10::1	2001:db8:10::1	**240b:0010:0200::/40/41**
3			Prefix : 240b:0010:0280::/41 **AAAA : 2001:db8:10::2**	2001:db8:10::1	**240b:0010:0200::/40/41**
4		240b:0010:0300::/40/0	Prefix : 240b:0010:0300::/40 AAAA : 2001:db8:10::5	2001:db8:10::5	240b:0010:0300::/40/40
5		240b:0010:0400::/40/0	Prefix : 240b:0010:0400::/39 AAAA : 2001:db8:20::1	2001:db8:20::1	240b:0010:0400::/40/39
6		240b:0010:0500::/40/0		2001:db8:20::1	240b:0010:0400::/40/39

- No.1はScope Subnetが/0で返ってくると予想していたが、結果は/39で返ってきた
  - 次の設定済みPrefixが240b:0010:0200::/40から始まるので、その中で一番大きいsubnet maskを返答している
  - 権威DNSサーバーの実装によっては/0が応答される場合もあり、実装による差異が見られる
  - Subnetを重複させるとキャッシュDNSサーバー側のキャッシュが上書きされてしまう場合があるので注意が必要
- No.2,3は権威DNSサーバー側で/41と設定しても、キャッシュDNSサーバー側のprefix設定が/40となっているので、キャッシュDNSサーバーからのリクエストが/40 でしか流れてこない
  →権威DNSサーバー側で/41に個別に設定しても回答が使われない(2001:db8:10::2が使われない)