

JANOG56 in Matsue Day1 2025年7月30日(水) 16:00～17:00

ECS(RFC 7871)を实际使おうと思うと困ること

～商用利用を前提とした検証結果の共有～ ※フルリゾルバ編

KDDI株式会社



自己紹介

- ・ 名前：今泉 充司(IMAIZUMI Atsushi)
- ・ 所属：KDDI株式会社 / プラットフォームエンジニアリング部(PE部)
- ・ 業務：KDDIの主なフルサービスリゾルバ(キャッシュDNS)の設計・開発・運用
- ・ 今回の発表までの経緯：
 - 社内外複数からECSの導入状況についての問い合わせを受けていた
 - JPIX社よりJストリーム社含めた3社合同検証の話を持ち掛けていただく
 - 試験するなら発表しようということになる
 - Janog56にエントリー / DNS Summer Day 2025にエントリー
 - 6/27 DNS Summer Day 2025に登壇 (かなり内容重複してます)
 - 7/30 Janog56に登壇
- ・ 免責：本プレゼン内の記述は今回の検証にあたり発表者個人の見解を示すものであり、所属する組織の見解やソフトウェア開発元の見解を示すものではありません。

プラットフォーム
エンジニアリング部

ECS(EDNS Client Subnet)とは

ECS(EDNS Client Subnet)とは

EDNS Client Subnet は DNS問い合わせ元の情報伝達する技術

EDNS Client Subnetとは、DNSの拡張プロトコルであるEDNS0^{*1}を用いて、問い合わせ先のDNSサーバに対して問い合わせ元の情報伝達する技術であり、RFC7871^{*2}で標準化されています。

今日、同じ内容のDNS問い合わせに対し、ユーザーがネットワーク上のどの位置にいるかによって、回答を変化させる権威DNSサーバが多数存在します。その理由は、地理的・ネットワーク的にサービスを分散している組織が、ユーザーを適切な接続先へ誘導することによって、遅延の減少やレスポンスの向上などを実現するためです。これらのサーバは、主に問い合わせ元のIPアドレスによってその位置を判断していますが、ユーザー位置の把握をより正確に行うための技術として、EDNS Client Subnetが策定されました。

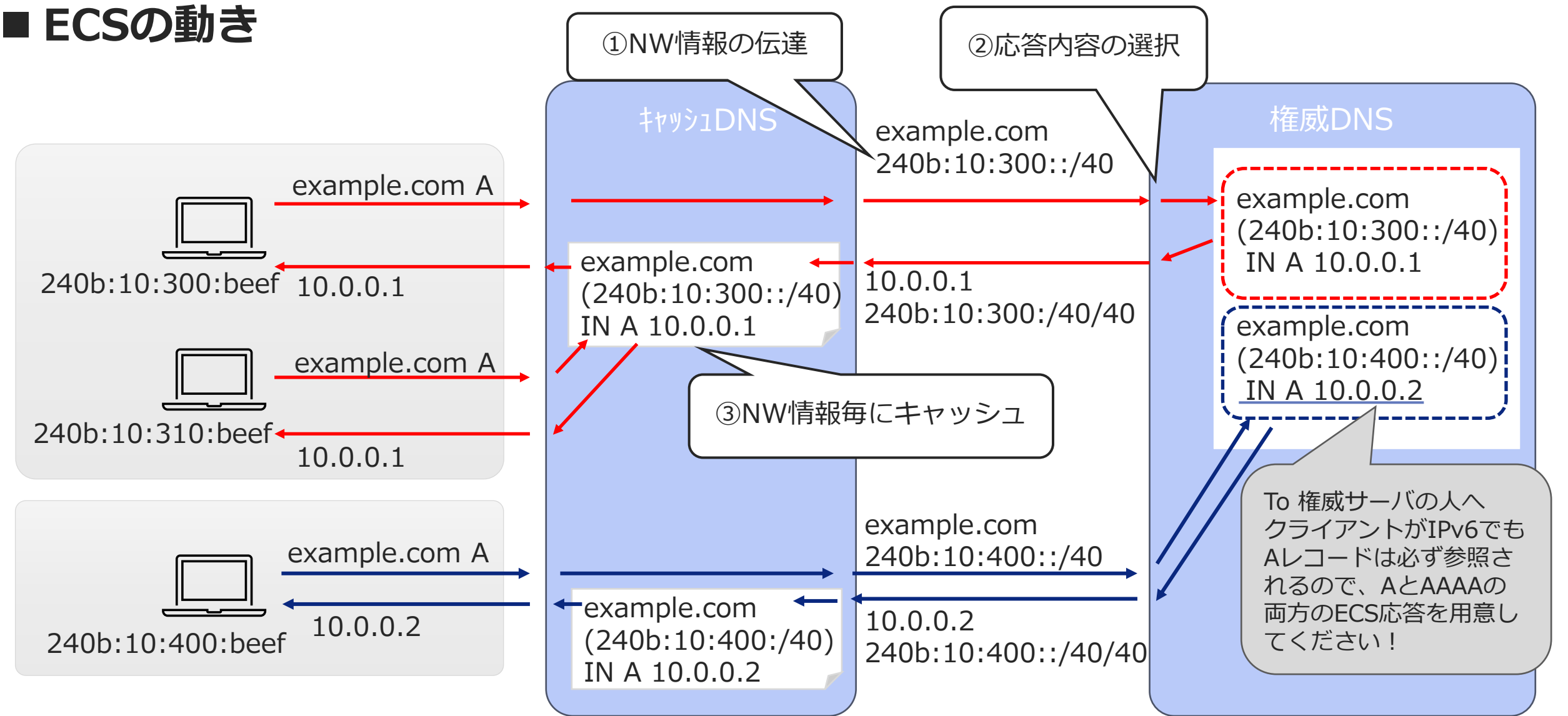
出典: [EDNS Client Subnetとは](#) (JPNIC)

■ ECSは、DNS問い合わせ元の情報(NW情報)を伝達し、応答する技術

- ① キャッシュDNSは、問合せ元のNW情報を権威サーバに伝達する
- ② 権威DNSは、伝達されたNW情報を元に事前定義された情報を応答する
- ③ キャッシュDNSは、権威DNSから得られた応答をNW情報毎にキャッシュする

キャッシュDNSから見たECSの動き

■ ECSの動き



キャッシュDNSからみたパケットキャプチャ

No.	Source	Destination	Prot	Client Subnet	Sour	Scop	Info
23	240b:10:300::1	キャッシュDNS 125:9::5	DNS				Standard query 0xfa5a AAAA te
24	キャッシュDNS 125:9::5	2401:dc40: 権威DNS :240:4	DNS	240b:10:300::	40	0	Standard query 0xe1cc AAAA te
25	2401:dc40 権威DNS :240:4	キャッシュDNS 125:9::5	DNS	240b:10:300::	40	40	Standard query response 0xe1c
26	キャッシュ 125:9::5	240b:10:300::1	DNS				Standard query response 0xfa5

ここに注目

• ECSパケットキャプチャ

- Stab(クライアント)はただのDNSクエリを投げている
- キャッシュDNSからは Subnet/Source Netmask(Prefix長)をリクエスト
- 権威からは Scope Netmask(Prefix長)を応答
- キャッシュDNSはStabにAnswerのみを応答

```
Option: CSUBNET - Client subnet
Option Code: CSUBNET - Client subnet (8)
Option Length: 9
Option Data: 00022828240b001003
Family: IPv6 (2)
Source Netmask: 40
Scope Netmask: 40
Client Subnet: 240b:10:300::
```

キャッシュDNSのDNSキャッシュダンプ

キャッシュDNSのキャッシュダンプ

```
~~~
; answer
verify-stream.com.      230      NS      dns-v17.verify-stream.com.
; pending-answer
dns-v17.verify-stream.com. 145 A      XXX.XXX.240.4
; pending-answer
dns-v17.verify-stream.com. 145 AAAA  2401:dc40:XXXX:XXXX:XXXX:XXXX:240:4
~~~
; answer
;test1.verify-stream.com. 419  A      YYY.YYY.51.61 ;; address prefix = 240b:10:300::/40
; answer
;test1.verify-stream.com. 425  AAAA    2001:260:YYYY:YYYY:YYYY:YYYY:51:61 ;; address prefix = 240b:10:300::/40
; answer
;test1.verify-stream.com. 486  AAAA    2001:260:YYYY:YYYY:YYYY:YYYY:51:60 ;; address prefix = 240b:10:200::/40/41
;
; Address database dump
~~~
```

ここに注目

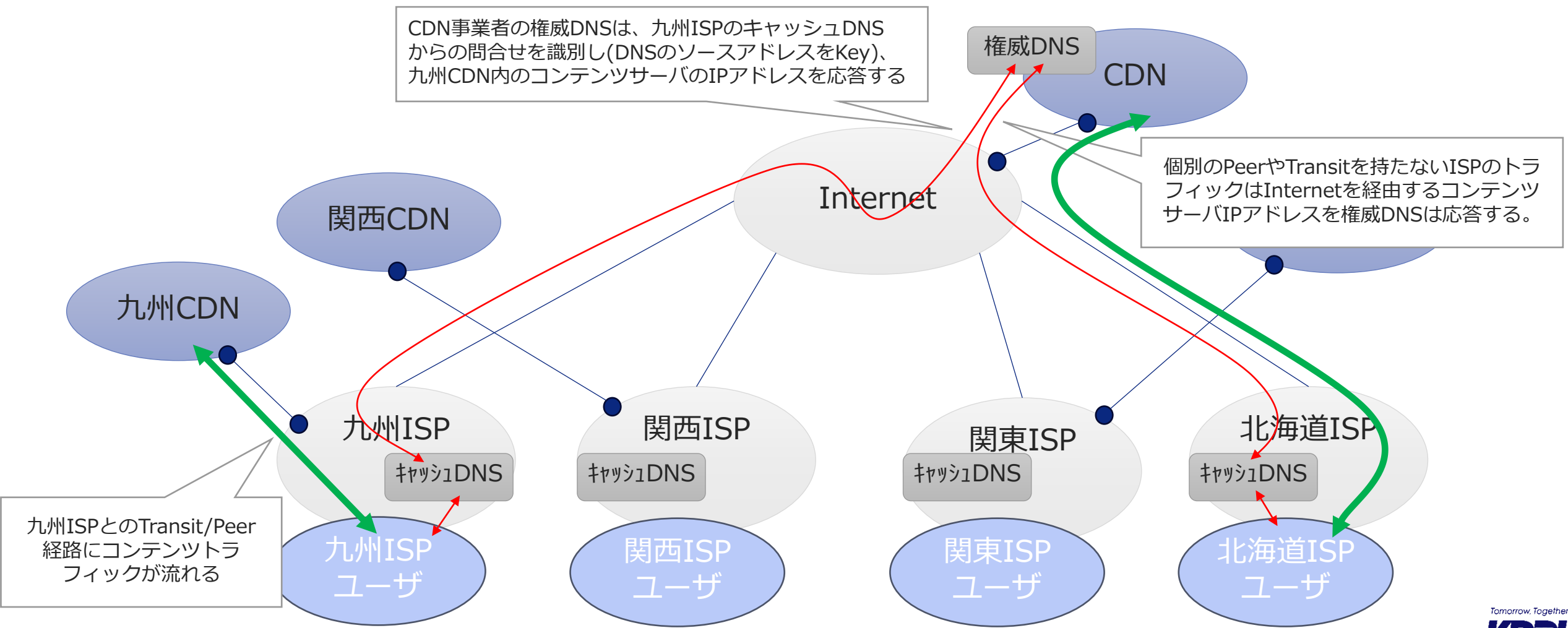
■ キャッシュの状態

- address prefix情報が付与された状態でキャッシュされている
- 同じレコードが 異なる address prefixでキャッシュされている

ECS導入シーン

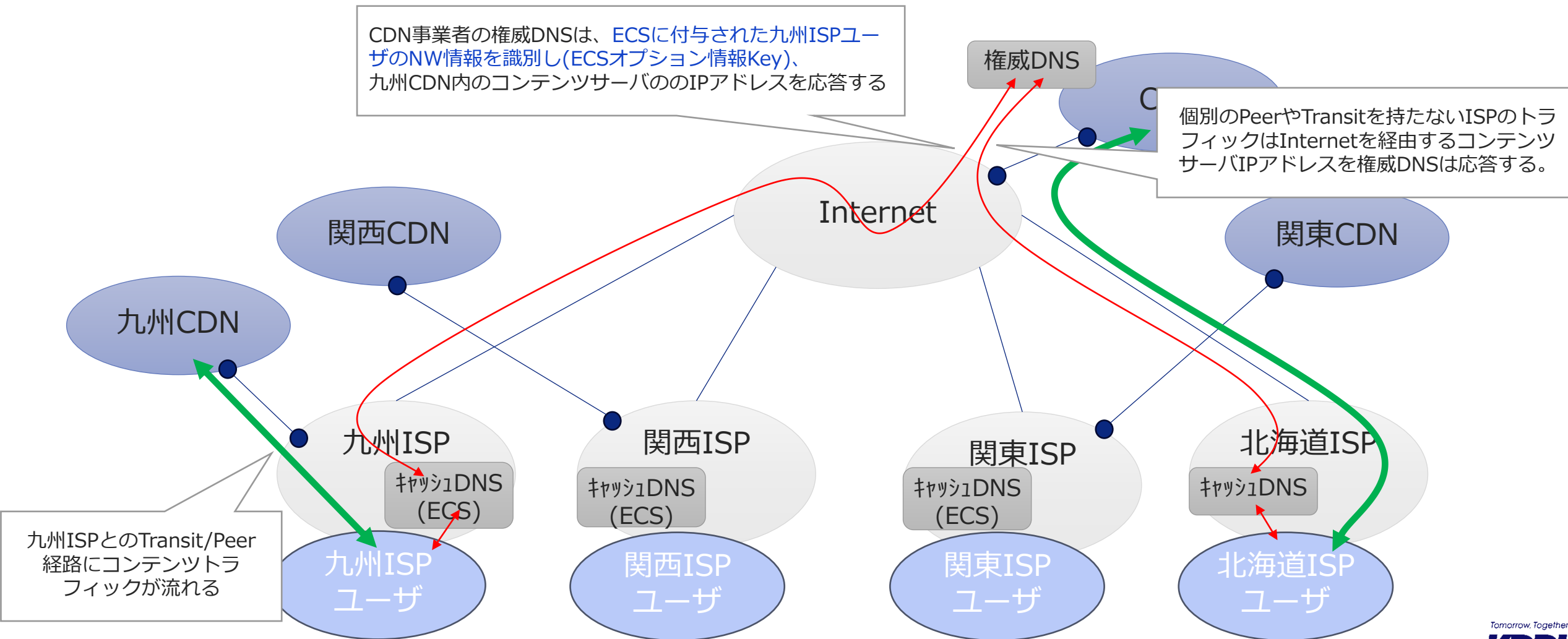
既存CDNコントロールの方式(DNS方式) ※ECSを使わない現行方式

- 各ISPのDNSサーバのIPアドレス情報を元に、CDNコンテンツサーバのIPアドレスをコントロールすることでコンテンツトラフィックコントロールを実現する(DNS方式)



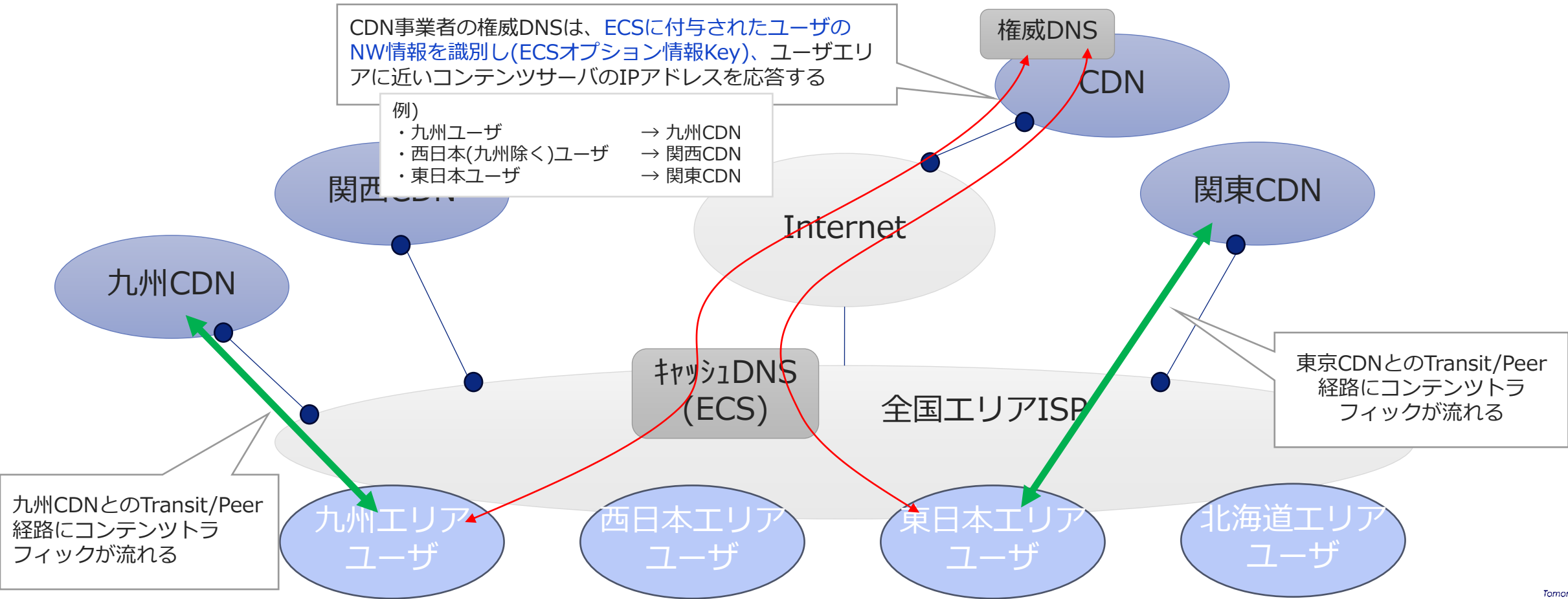
ECSを活用したCDNコントロールの方式シーン①(ISPの選択)

- ECSで付与されたユーザのIP情報を元に、CDNコンテンツサーバのIPアドレスをコントロールすることでコンテンツトラフィックコントロールを実現する (既存DNS方式と変化がない)



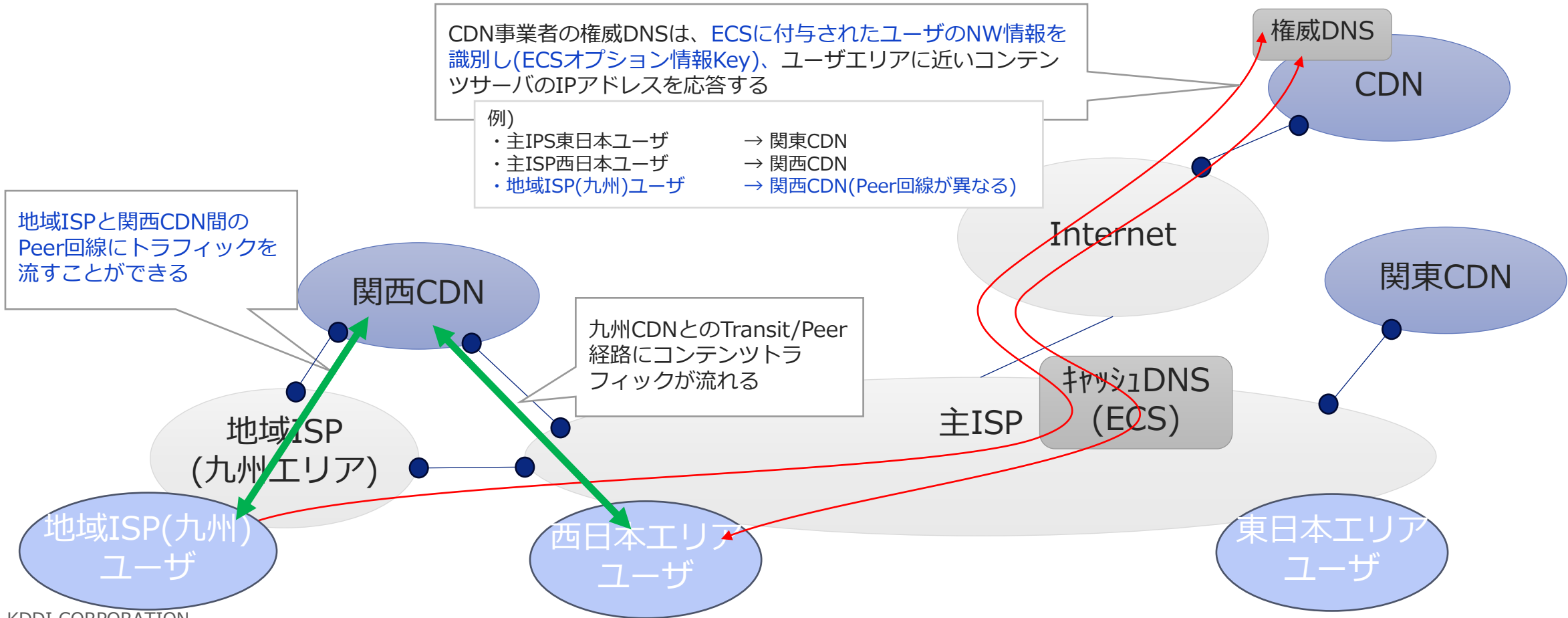
ECSを活用したCDNコントロールの方式シーン②(エリアの選択)

- ECSで付与されたユーザのIP情報を元にCDNコンテンツサーバのIPアドレスをコントロールすることで、全国エリア展開をしているISP網内でトラフィックコントロール
この方式ではISP網内に複数のTransit/Peerがある場合にエリア毎のコントロールが可能に



ECSを活用したCDNコントロールの方式シーン③(DNSの外部提供)

- ECSで付与されたユーザのIP情報を元にCDNコンテンツサーバのIPアドレスをコントロールすることで、コンテンツトラフィックを地域ISP向けのPeer回線に流すことが可能。
主ISPのキャッシュDNSを利用している地域ISP網内でもトラフィックコントロールが可能

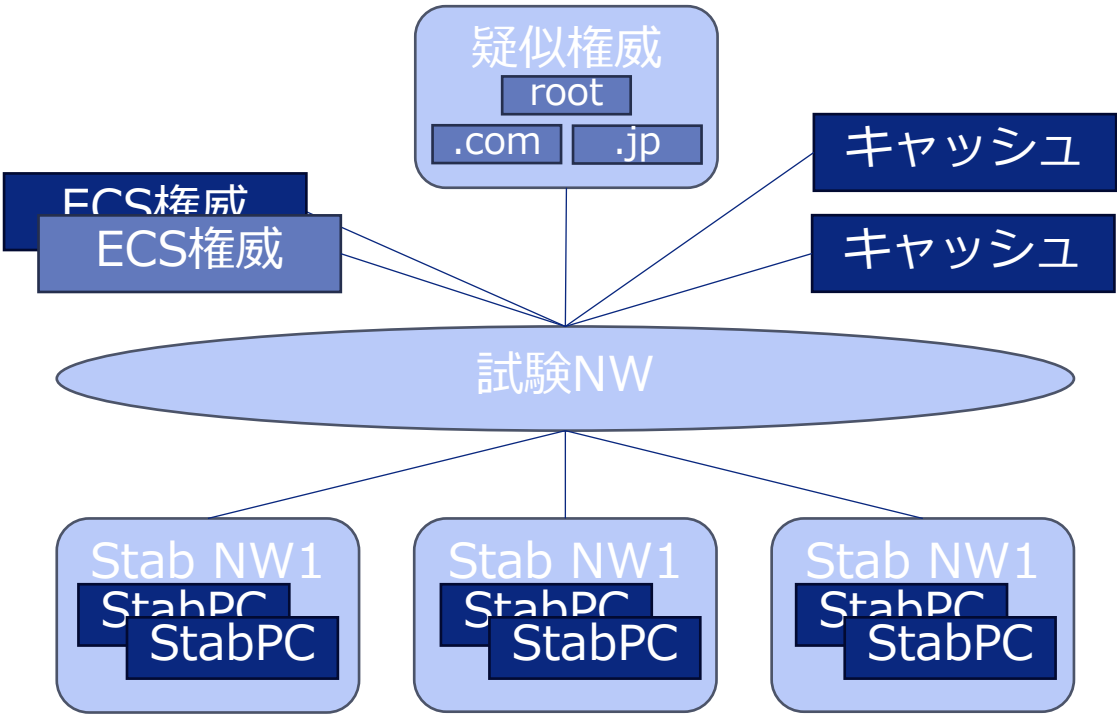


ここから試験結果の紹介

今回の試験環境の構成

試験NW

■ 試験はスタンドアロン環境

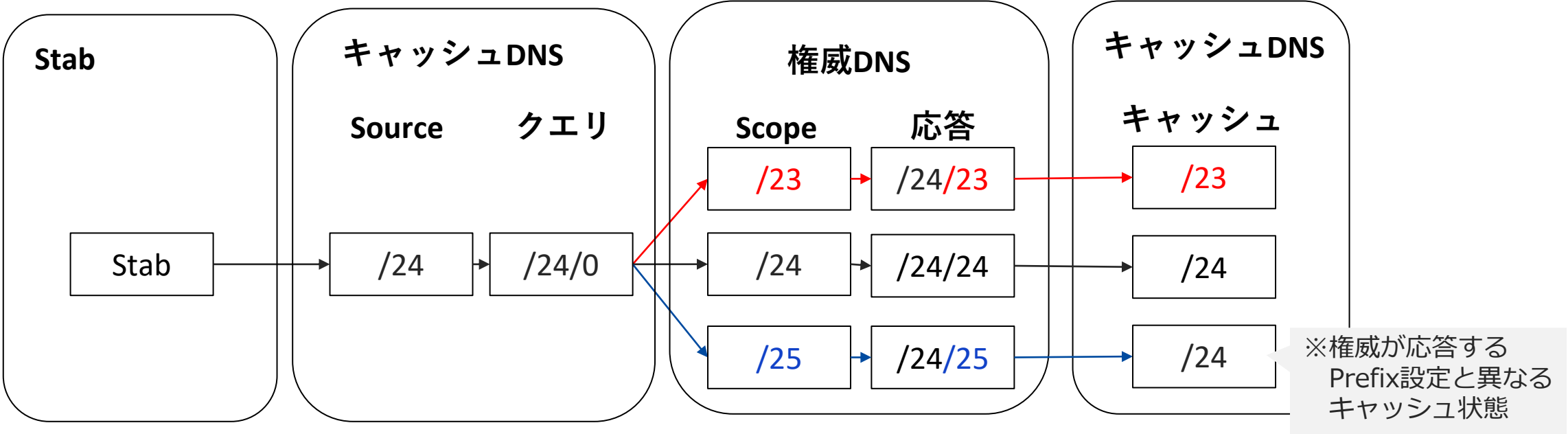


構成機器

- ECSキャッシュDNSサーバ：2台
 - 製品A：A社ソフトウェア
 - 製品B：OSS Bベースアプライアンス
- ECS権威DNSサーバ：1台
 - ソフトウェア F
 - ソフトウェア K(動作確認用)
- 権威DNSサーバ：1式
 - 疑似root/疑似.com/疑似.jp
- その他設備
 - NW機器類
 - PC(Stab/Capture等)
 - etc

ECS基本動作部分比較

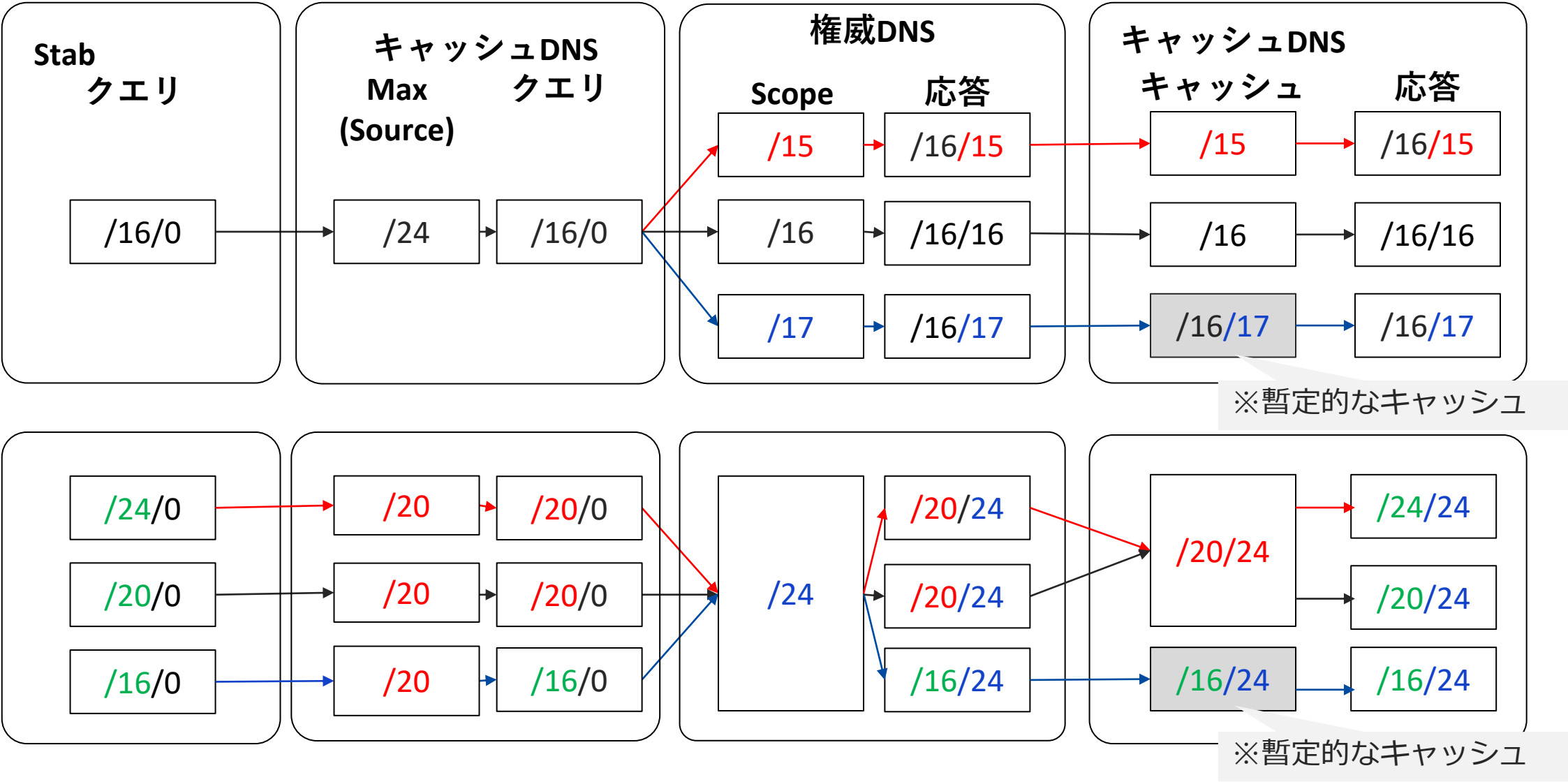
■ 基本的な動作に差分無し



- 2種ソフトウェアで異なった動作箇所
 - RRの種類によるECSオプション付与に違い
 - A製品：すべてのRRにECSオプションを付与する
 - B製品：特定のRRにはECSオプションを付与しない

付与しないRR	DS	NS	SOA	DNSKEY	NSEC	NSEC3					
付与するRR	A	AAAA	CNAME	TXT	PTR	MX	HTTPS	SPF	SVR	NPTR	ANY

クライアントECS在り(ECSフォワード有効)



※暫定的なキャッシュ

※暫定的なキャッシュ

その他動作の異なった個所(一部)

■ キャッシュした情報の表示形式が異なる

A製品

address prefix 240b:YYYY:XXXX::0/40

address prefix 240b:YYYY:XXXX::0/39

B製品

address prefix 240b:YYYY:XXXX::0/40

address prefix 240b:YYYY:XXXX::0/39/40

■ ECSフォワード設定時のstabからのプライベートアドレス処理に動作の違い

A製品：プライベートアドレスを+subnetで付与した問い合わせはすべてRefuse

B製品：プライベートアドレスを+subnetで付与した問い合わせもすべてForward

■ ECSフォワード非設定時の stabからの+subnet付与処理にA製品だけ例外あり

基本的には一律Refuse応答

A製品のみ：Stab Subnetと +subnetの範囲が一致した場合は問い合わせされる

課題

リソースに関する課題

- ① Subnetの単位でキャッシュするということは、同じFQDNに対するエントリ数は最大で、
『キャッシュDNSの参照を許可するNWサイズ ÷ Source Prefixサイズ』 倍に増加する

例) $106.128.0.0/10 \div /24 = 16,384$ 倍

$106.128.0.0/10 \div /16 = 64$ 倍

$240F::/23 \div /40 = 131,072$ 倍

$240F::/23 \div /32 = 512$ 倍

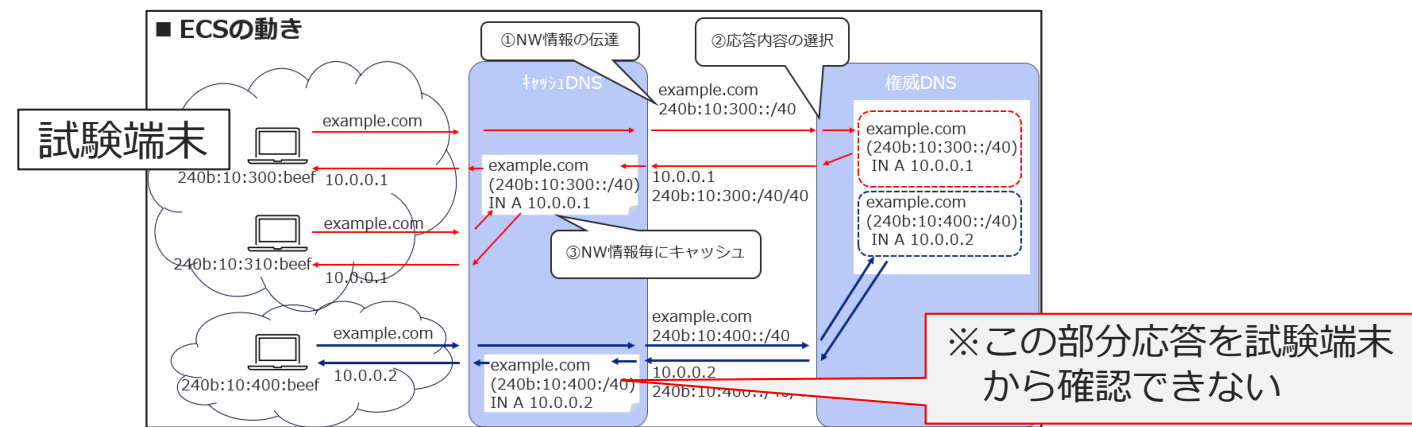
⇒キャッシュメモリ数の増加の懸念

- ② 同様の理由により、非再帰問い合わせ数も増加する

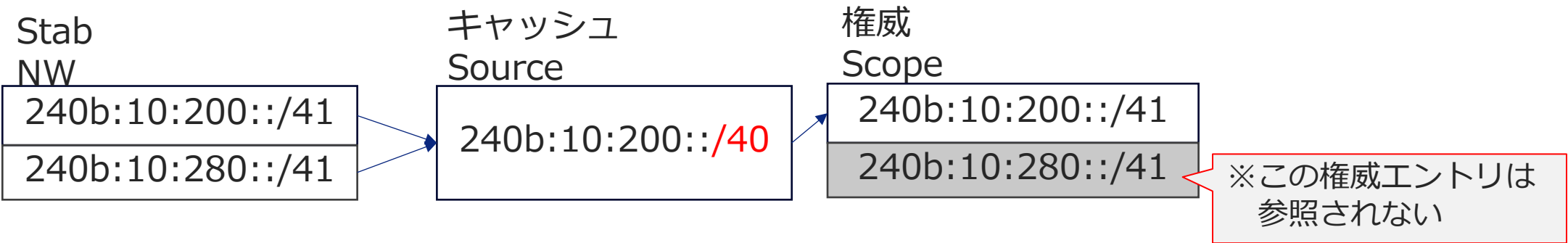
⇒キャッシュHIT率の低下の懸念

運用面による課題

- ③ ECSの応答状態、キャッシュ状態をDigコマンドで確認することが困難
普通にDigを投げただけだと試験端末のサブネットの応答しか確認できない。



- ④ Source Prefix長 < Scope Prefix長 の設定を行うと、権威側で想定していない (参照されない権威エントリ) 応答になってしまう。



導入に向けた課題

■リソース課題(キャッシュ肥大化/キャッシュHIT率低下対策)

- リソース対策の出来るアプリケーションの利用やシステム設計が必要
例) ・ 複数の連続しないサブネットをまとめる機能を有するDNSソフトウェアを利用する
・ ECS対応ドメインが少ないうちはECS対象ドメインのキャッシュTTLを短縮し、キャッシュ領域の肥大化を抑制する
・ etc.

■運用課題

- DNS応答結果を確認する環境を用意する
例) ・ ECSエリア毎の試験サーバを準備する
・ 限定的なNWからのみ、クライアントからのECSオプション付与を許容する設定を行う
- ECSに関するリソース監視を行う
例) ・ ECSクエリを含むクエリ数を別統計で集計する
・ クエリLogにECSオプション情報を出力し、クエリ統計を可視化する
・ etc.

■権威DNS連携課題

- キャッシュDNSと権威DNSの動作を確認できる環境を用意する(FT環境?)

「つなぐチカラ」を進化させ、
誰もが思いを実現できる社会をつくる。

KDDI VISION 2030

