

CATVネットワークにおける

DHCPオペレーション

～開発の立場編～

2019年7月26日

株式会社XACK

矢島 崇史



# 自己紹介

- 矢島 崇史
- 株式会社XACK
  - RADIUS
  - DHCP
  - DNS

などのUNIX系OSで動作するサーバーアプリケーションを作っています。

# DHCP開発経緯

- あるキャリアでDHCPv6が必要となった
- 当時、キャリア向けに使用できるDHCPv6サーバーがCNRしかなかった(らしい)
- DHCPv6サーバー作ってよ
  - 作った

# DHCP開発経緯

- せっかく作ったから他キャリアにも売りたいくなった
  - 「DHCPv4が欲しいんだけど」といわれる
- 作った
  - 誤 御社製品はDHCPv6に対応しているでしょうか
  - 正 弊社製品はDHCPv4にも対応しています
- 「DOCSISはTFTPも扱えないと」といわれる
  - 作った

# キャリアの要件 (冗長)

## ■ 3台以上で冗長したい

- IPの払い出しが止まると全サービスが止まるといっ  
て過言でない

- 数が並んでるだけでも安心できる

- 何台構成でもOK!

## ■ DRサイト、別局舎で冗長したい

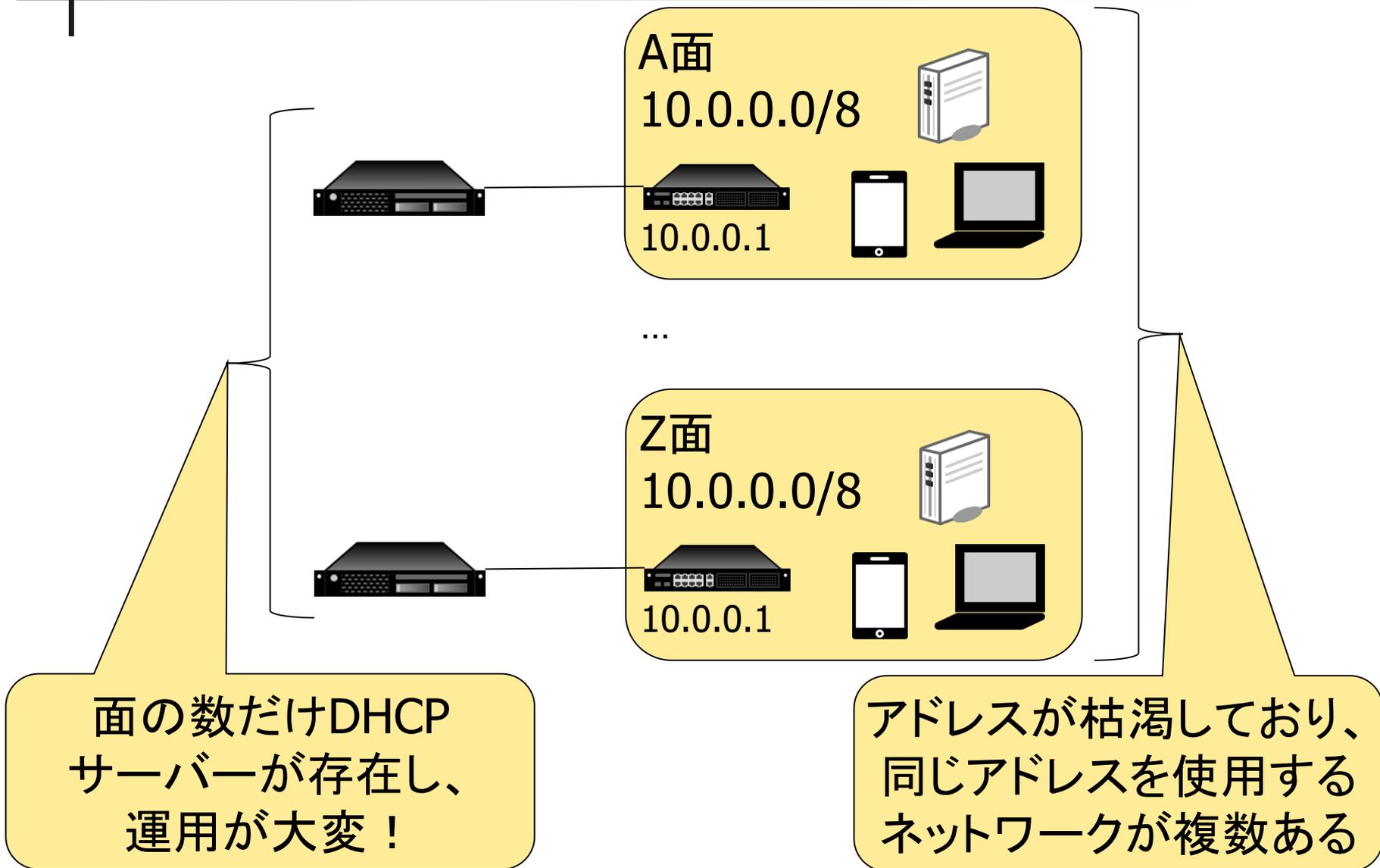
- 東京が壊れても大阪では動いてほしい

- でもIPアドレスは有限

- 別局舎でも同期可能 (遅延はあるけれど)

- リアルタイムにホットスタンバイで同期するよりはバックグ  
ラウンドでコールドスタンバイの方が結局はおすすめ

# キャリアの要件(サーバー集約)



# キャリアの要件(サーバー集約)



A面

10.0.0.0/8



10.0.0.1



...

Z面

10.0.0.0/8



10.0.0.1



通信するIFで面を判別、  
OSのポリシールーティン  
グで同一アドレス宛の  
パケットを正しい面に  
届けられるようにし、  
サーバーを集約

# キャリアの要件(多段リレー)

ルーティング集約、アドレス節約のためBの情報でプールを定めたい。でもgiaddrはAが入っている。クライアントに回答を届けるため書き換えられない。そのため、**L3ルーティングのついでにOption82を追加する**

リレーはエッジ寄りのL3SWが実施する。リレー先はDHCPサーバー



DHCPサーバーは**Option82でアドレス決定**、giaddr宛に応答する。

クライアントはブロードキャスト

- DHCPはリース情報を(ディスクなどに)永続的に保持することがプロトコルとして前提となっている
- OSSのDHCPサーバーはこの辺がボトルネックになって性能が出ないことがある
- Linuxでは、極短時間に同一ファイルに多数ランダム書き込みをするとI/Oが停止する
  - 仕方ないのでいい感じの独自のファイルDBを作成する羽目に

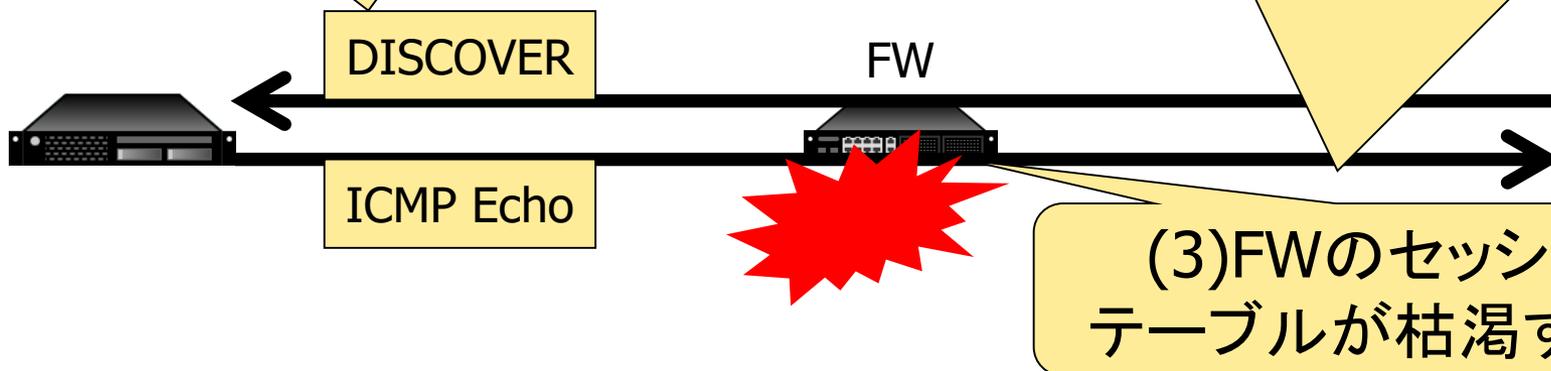
# 開発で困ったこと(ICMP Echo問題)

DHCPv4では払い出すアドレスが使われていないか確認しないといけない(SHOULD)

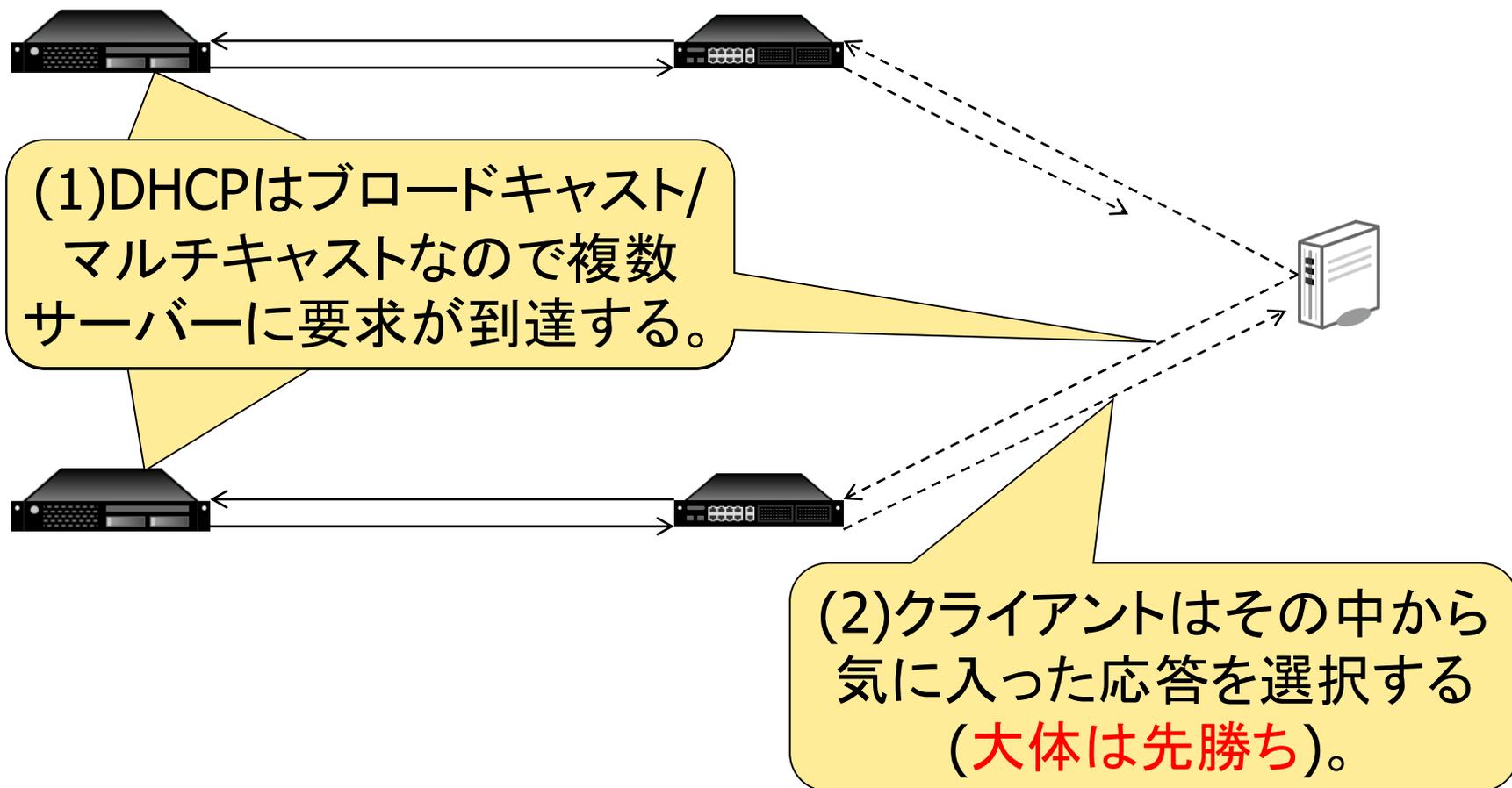


(1)バーストラフィックが発生すると.....

(2)Pingもバーストするので.....

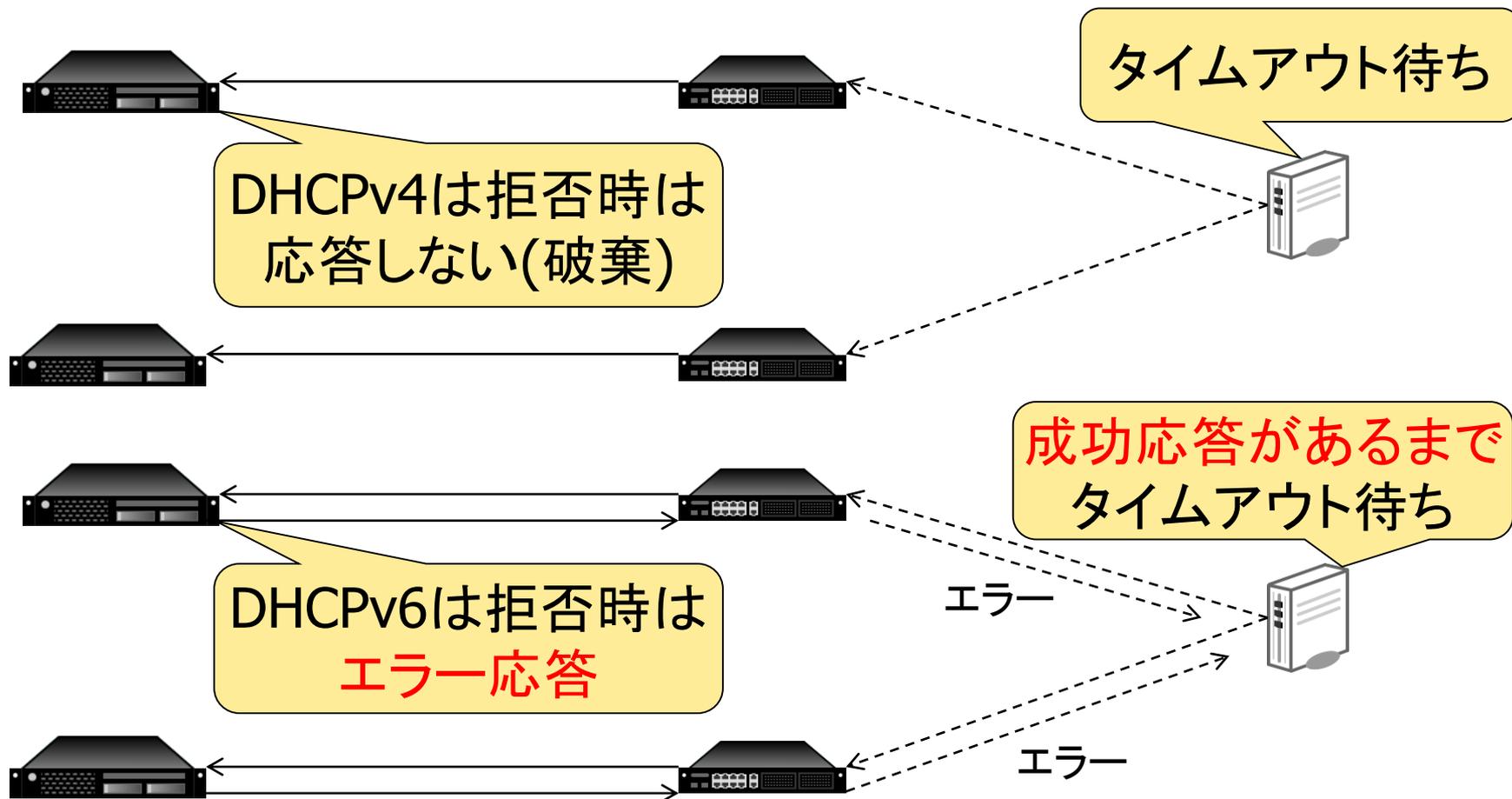


# 開発で困ったこと(リトライ問題)



プロトコルベースの冗長性

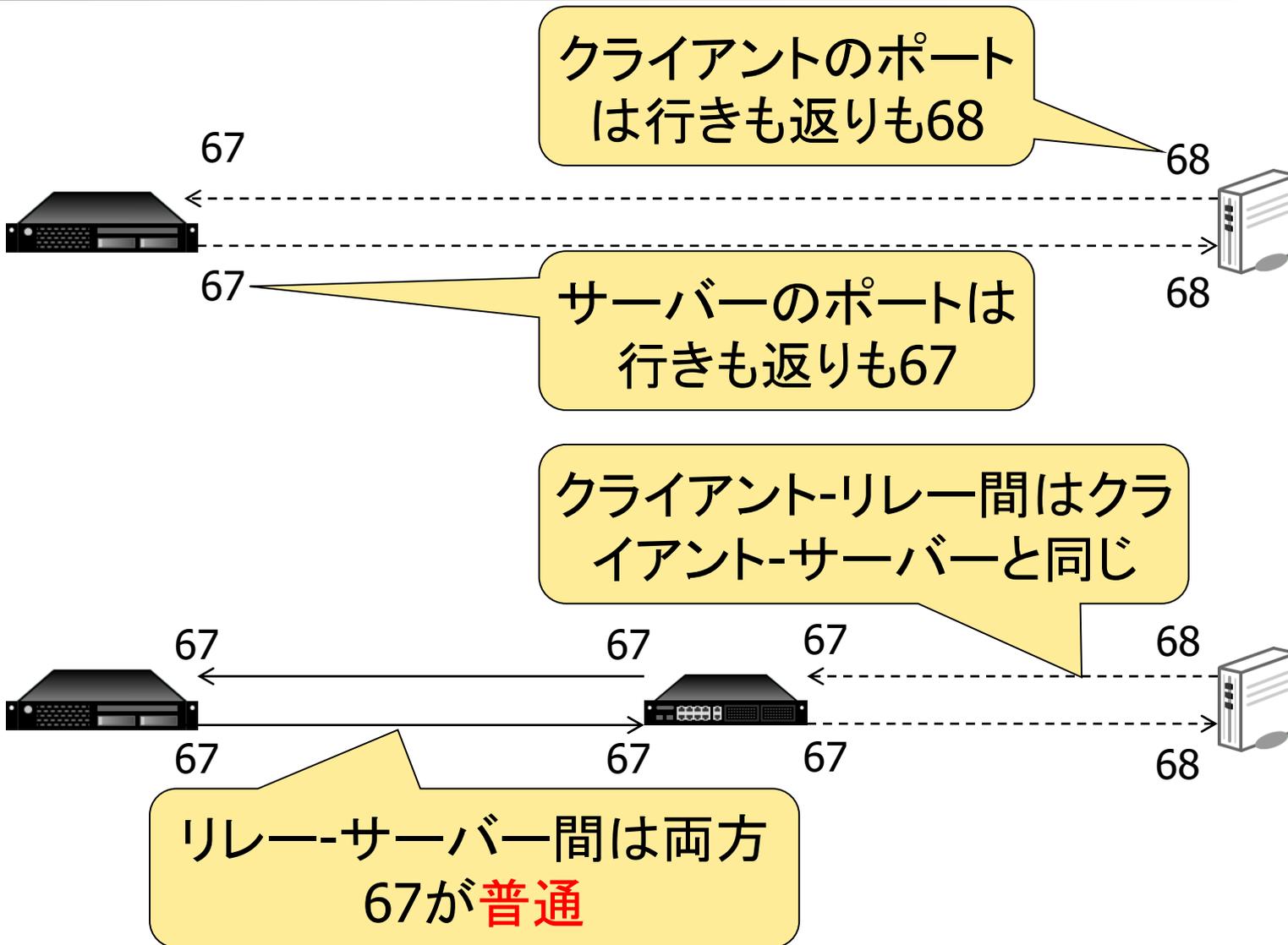
# 開発で困ったこと(リトライ問題)



DHCPv4/v6の動作の違い

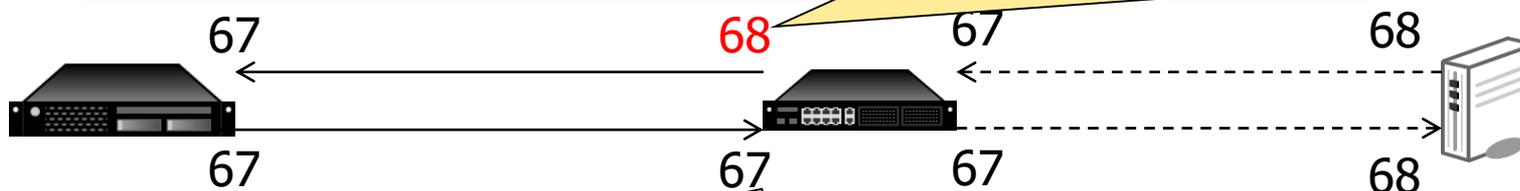


# 開発で困ったこと(ポート番号問題)



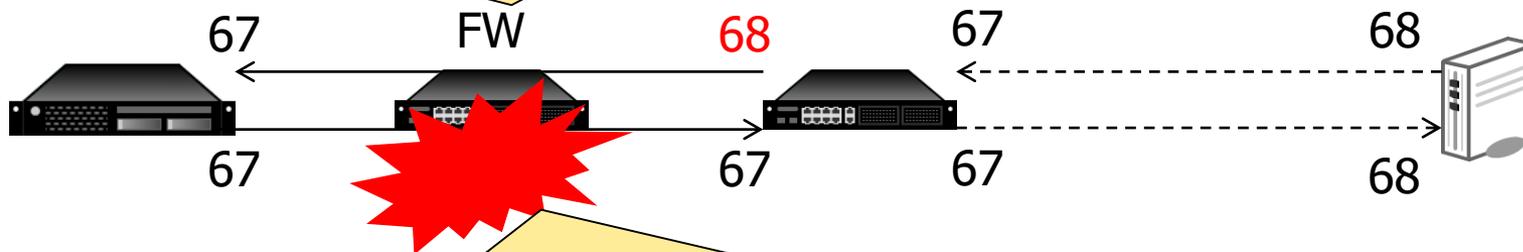
# 開発で困ったこと(ポート番号問題)

(1) 送り元を68にするリレーエージェントが存在する(サーバーから見たらクライアントだから?)。



(2) でも応答は67で受け取る。

(3) FWは大抵アドレス/ポートでセッションを識別するので、行きと帰りが同じセッションだと認識できない



(4) FWのセッションテーブルが枯渇する!

# 振り返ると

- ネットワークの負荷で苦しめられたことが多かった様子
  - FWのセッションテーブル、ブロードキャスト/マルチキャストの輻輳
- これが顕在化するにはトラフィックが大きく起因するはず
- いくつかのキャリアのトラフィックを比較してみました

# トラフィックモデル

	A社 (CATV)	B社 (CATV)	C社 (CATV)	D社 (CATV)	E社 (FTTH)	F社 (WiFi)	XACK DHCP (参考)
DISCOVER (件/日)	98,000	6,900	12,000	80,000	36,000,000	2,600,000	430,000,000
OFFER (件/日)	95,000	6,900	12,000	80,000		2,600,000	
REQUEST (件/日)	880,000	58,000	170,000	450,000		3,800,000	
ACK (件/日)	880,000	60,000	170,000	470,000		2,500,000	
端末	46,000	2,500	5,200	14,000	1,700,000	500,000	
リース時間	12時間	2/9時間	2/9時間	2/24時間	24時間	10分	

# 比較の考察

- 移動系と固定系はモデルが全く違う
  - DISCOVER/REQUEST比、リース時間
  - REQUESTにACKが返らないのはAPを移動している(違うサブネットに移動している)から
- FTTHとCATVは似ている様子
- CATVは端末数少なめ
  - この規模だと十分耐えられる負荷に収まるのでは
  - バースト時はモデルが違うのでもしかしたら.....
    - 広域停電からの復旧
    - 端末アップデートやスリープ復帰など特定時刻イベント



<https://www.xack.co.jp>