CATVネットワークにおける DHCPオペレーション ～運用の立場編～

2019年 7月26日
janog44@神戸
ケーブルテレビ株式会社
技術部　日里友幸
<nissato@cabletv.jp>

会社紹介

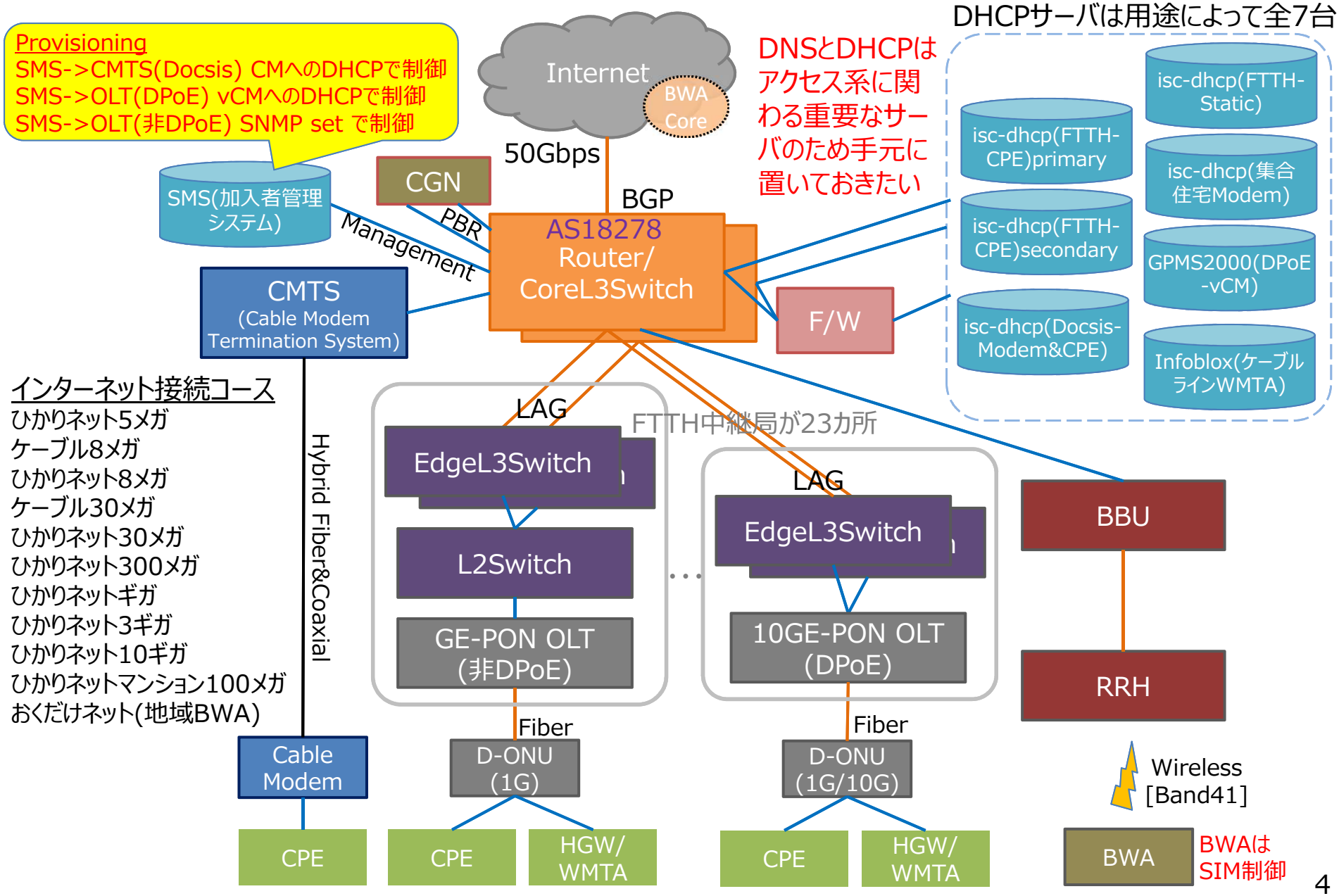
- 栃木県栃木市に本社のあるケーブルテレビ会社。
- 栃木県・群馬県・茨城県の3県にまたがって、5市5町にエリア展開。いち早く12年前からFTTH導入。
- インターネットのユーザ数は約5万件(うち6割がギガへ加入)、2018年からは10ギガサービスも開始。
- AS18278、上位トラフィックはピークで約25Gpbs、大阪にも直接接続しています:D

蔵の街とちぎの夜景(小江戸・小京都とも呼ばれる)

自己紹介

- 氏名 日里 友幸 (ニッサト トモユキ)
- 1980年栃木県うまれ、栃木県育ち
- 趣味 カメラ、ゴルフ、神社仏閣・古墳城郭巡り
- janog歴　janog31六本木、janog35静岡、janog40郡山、janog44神戸で4回目
- ケーブルテレビに入社して16年、主にIPネットワークの構築・設計・保守・運用・管理に従事、管理的には放送設備にも関わる、最近では4K放送、地域BWA、ローカル5G、LPWA、等にも携わる。

自ISPネットワーク構成概要



DocsisとDPoE(CATV特有の規格)

➢ DOCSIS(HFC)

Data Over Cable Service Interface Specifications の略で、米国CableLabsが策定する、ケーブルテレビ回線を利用して高速なデータ通信を行うための規格。
最新のDOCSIS3.1では高次変調(512~4096QAM)とChannel Bondingや誤り訂正等により下り最大10Gbpsのスループットを実現。
CableModemのMACアドレスとDocsisConfigファイルをDHCPで紐付けてDHCP割当後にtftpでダウンロードさせてQoS等を実現している。

➢ DPoE(FTTH)

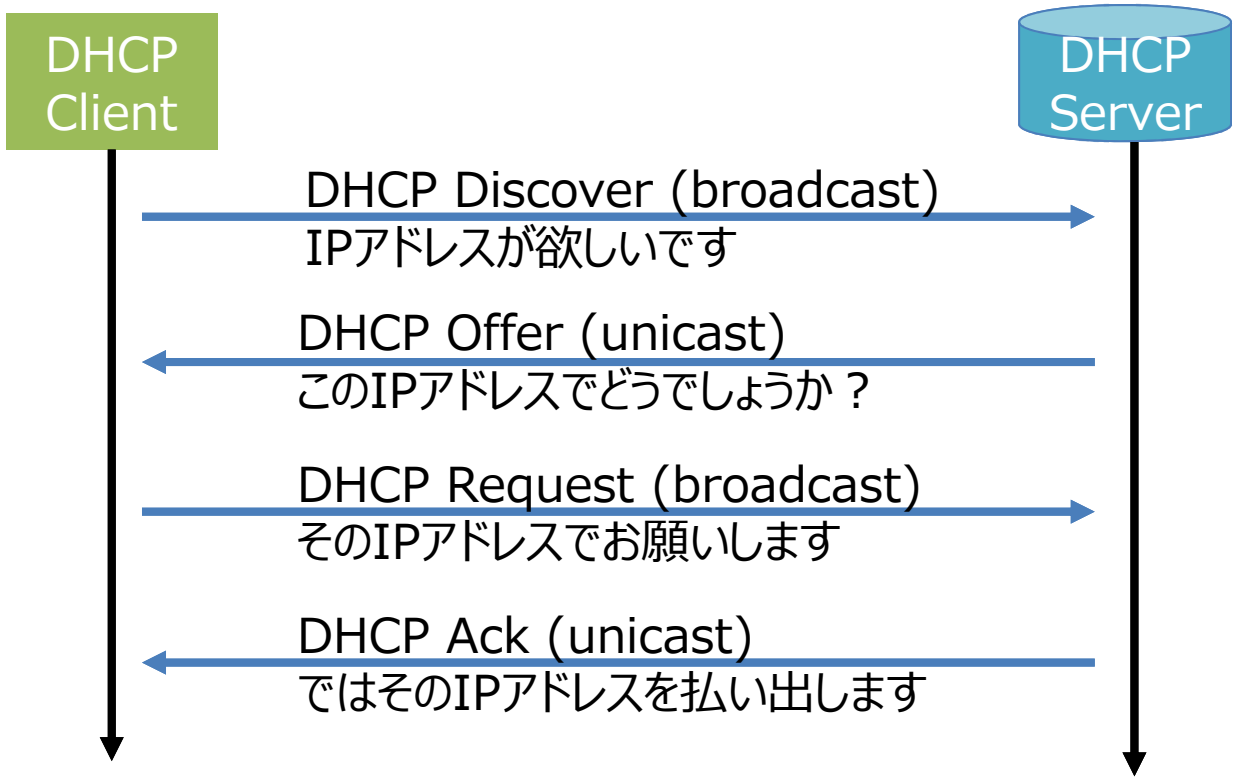
DOCSIS Provisioning of EPON の略で、米国CableLabsが策定するDOCSISとEPONの統合ネットワーク規格。
DOCSISで使用していた資産(DHCPサーバ等)や連携APIをEPONネットワークでも引き継げるメリットがある。
D-ONUをvCM(virtualCableModem)と見立て、CableModemと変わらない制御を実現している。
※当社では先行した非DPoEのFTTHインフラはSNMPで制御している。

DHCPサーバのいろいろ

- isc-dhcp
- CNR(Cisco Network Registrar)
- PNR(Cisco Prime Network Registrar)
- isc-KEA
- Microsoft Windows DHCP
- アプライアンス(isc-dhcp派生、フルスクラッチ等)
- ブロードバンドルータ

※今回の設定例はisc-dhcpを基に事例紹介します。
※行数省略のためインデントが変な箇所がありますが気にしないでください。

DHCP(IPv4)基本シーケンス



あとこの辺りを覚えておくとdhcpd.log解析で便利

- lease-timeの1/2でDHCP Request(unicast)更新
- lease-timeの7/8でDHCP Request(broadcast)更新 ※上記DHCP Request(unicast)が失敗した場合
- dhcpd.leasesのtime-stampはGMT表記

単一サブネット

- DHCP Discoverはbroadcastなので、届く範囲はbroadcastドメイン内のみ。
- つまりL2内しか伝搬しないので、L3を越えられない。

dhcpd.conf

```
subnet 192.168.0.0 netmask 255.255.255.0 {
        option broadcast-address    192.168.0.255;
        option subnet-mask          255.255.255.0;
        option routers              192.168.0.1;
        range                       192.168.0.2 192.168.0.254;
}
```

- けれども、L2ネットワーク毎にDHCPサーバをたてるのは大変ですよね？→それを回避するのが次々頁の”DHCP relay-agent”

複数サブネットと shared-network

- 1つのsubnetではIPアドレスが不足した場合に、1つのL2ネットワーク内で複数subnetをアサインしたい場合は、shared-network設定を使う。
- 上位スイッチでは同一インターフェイにsecondaryでIPアドレスを追加。
- グローバルIPv4アドレスでは必須な運用かなと。

dhcpd.conf

```
shared-network hogemachi {
       subnet 192.168.0.0 netmask 255.255.255.0 {
                option routers 192.168.0.1;
                option subnet-mask 255.255.255.0;
                pool {  range 192.168.0.2 192.168.0.254;
                        option domain-name-servers 8.8.8.8,8.8.4.4;
                } }
       subnet 192.168.1.0 netmask 255.255.255.0 {
                option routers 192.168.1.1;
                option subnet-mask 255.255.255.0;
                pool {  range 192.168.1.2 192.168.1.254;
                        option domain-name-servers 8.8.4.4,8.8.8.8;
                } } }
```

DHCP relay-agent と GIADDR

- DHCP relay-agent機能により、DHCPサーバとDHCPクライアントが異なるsubnetに存在してもDHCPクライアントから受信したbroadcastをunicastに変換して、DHCPサーバに転送します。
- DHCPサーバはGIADDRを見てDHCPクライアントの所属するsubnetを識別します。

```
ip routing
int vlan hogehoge
  ip address 192.168.1.1 255.255.255.0 secondary
  ip address 192.168.0.1 255.255.255.0
  ip helper-address 10.0.0.100
int giga 0/1
  switchport access vlan hogehoge
```

primaryのGatewayアドレスがGIADDR

GIADDRを含むshared-networkからIPアドレスを払い出し

broadcastはL3を超えられないが、relay-agentによりunicast変換されてDHCPサーバまで届く

ここからは・・・、

DHCPの運用事例と
トラブル事例

DHCP message format

- DHCPには可変長のOption-Fieldがあります。

Message Type [1bit]	Hardware Type [1bit]	Hardware Address Length [1bit]	Hops [1bit]
Transaction ID [4bit]			
Second Elapsed [2bit]		Boot Flag [2bit]	
Client IP Address (CIADDR) [4bit]			
You(Client) IP Address(YIADDR) [4bit]			
Next Server IP Address (SIADDR) [4bit]			
Relay Agent IP Address (GIADDR) [4bit]			
Client MAC Address (CHADDR) [16bit]			
Server Host Name (SNAME) [64bit]			
Boot File Name (FILE) [128bit]			
Option [variable]			

次の運用事例で取り上げるDHCP Optionは以下。

- Option82; Relay agent information
- Option60; Vendor class identifier

運用事例1 DHCPによるUserTrace

- Option82を使って、IPアドレスからONUのMACアドレス(収容OLTのIFやIDも分かる)を追跡。
- 各L3スイッチで ip dhcp snooping 設定が必要。
- ONU上部のOLTで Option82 Enable が必要。
- IPソースガード(Static成りすまし防止)も推奨。

dhcpd.conf

```
if exists agent.remote-id {
        log (info, concat( "OPT82 IP Address ", binary-to-ascii(10, 8, ".",leased-address),
" RID: ", binary-to-ascii(16, 8, ".",option agent.remote-id)));
}
```

dhcpd.log

```
May 26 04:16:33 dhcp-ha1 dhcpd: DHCPDISCOVER from [CPE-MAC] via [GIADDR]
May 26 04:16:33 dhcp-ha1 dhcpd: DHCPOFFER on [IPv4-address] to [CPE-MAC] via [GIADDR]
May 26 04:16:33 dhcp-ha1 dhcpd: OPT82 IP Address [IPv4-address] RID:
2:2:0:[0xif]:3:8:0:[0xid]:[ONU-MAC;hd]
May 26 04:16:33 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 26 04:16:33 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
```

運用事例2 DHCPによるStatic割当

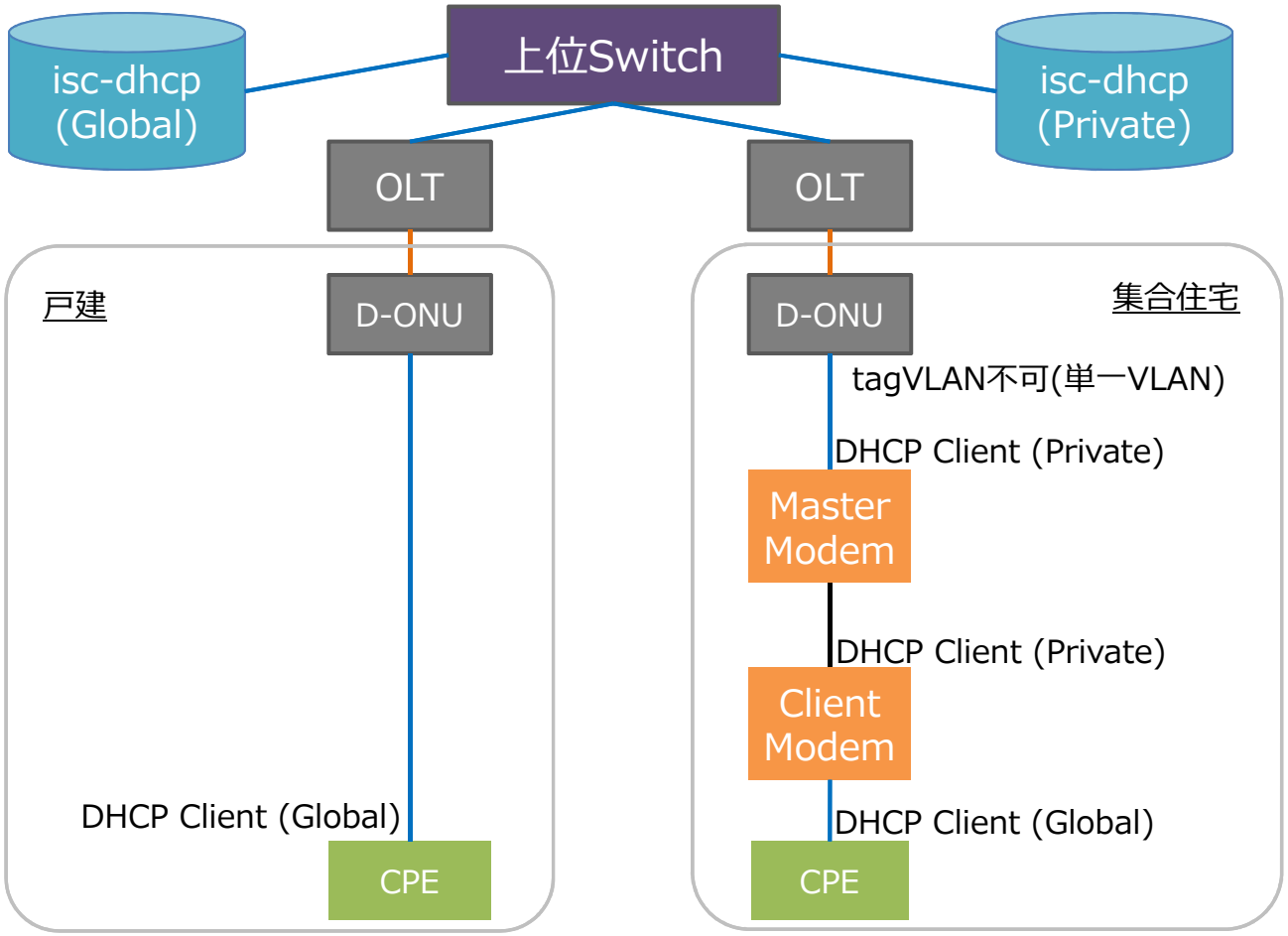
- Option82を使って、ユーザONUのMACアドレスとIPアドレスを静的に紐づけ、DHCPだけど動的ではない固定IP1個割当サービス提供。
- CPE-MACアドレスを紐づけるfixed-address運用もあるが、CPEはお客様都合で変わるため運用負荷を減らせる。けれどもCPEでDHCP Releaseがうまくいかないと結局対応が必要・・・。

dhcpd.conf

```
class  “hogestatic1” { match pick-first-value (option agent.remote-id); }
subclass  “hogestatic1” 2:2:0:[0xif]:3:8:0:[0xid]:[ONU-MAC];
shared-network static-ip1 {
     subnet 10.0.0.0 netmask 255.255.255.240 {
       option routers 10.0.0.1;
       option domain-name-servers 172.16.0.1,172.16.0.2;
       pool {
         allow members of "hogestatic1";
         range 10.0.0.2;
         }  }  }
```

運用事例3 DHCPサーバの並行稼働①

- 集合住宅でONU配下の単一VLAN内に、2つのネットワーク(Global/Private)、各々別のDHCPサーバからIPアドレスを割り当てる必要があった。



運用事例3 DHCPサーバの並行稼働②

- DHCP Discoverはbroadcastなので、IPネットワークではどちらかを意図的に選択させることはできないため、DHCPサーバが2つあればどちらにもDiscoverが届いてしまう。同一VLANからのDiscoverのためGIADDRも共通となる。
- 集合住宅モデムがDHCP Option60(vendor-class-identifier)に対応していたためclass定義し、dhcpd.confでallow/denyさせることで割り当てを制御できた。
- 割り当てをしない側のDHCPサーバでもpoolを持たないsubnetを記述しないと”no free leases”を発出してしまうので注意が必要。

運用事例3 DHCPサーバの並行稼働③

dhcpd.conf(Common)

```
class “hoge-master” {
   match if substring (option vendor-class-identifier, 0, 24) = “c.LINK-MODEM-Coordinator”;}
class “hoge-client” {
   match if substring (option vendor-class-identifier, 0, 19) = “c.LINK-MODEM-Client”;}
```

dhcpd.conf(Global)

```
shared-network hogekura{
       subnet A.B.0.0 netmask 255.255.255.0 {
                option routers A.B.0.1;
                option subnet-mask 255.255.255.0;
                pool {  deny members of  “hoge-master”;
                        deny members of  “hoge-client”;
                        range A.B.0.2 A.B.0.254;
                        option domain-name-servers 8.8.8.8,8.8.4.4; } }
        subnet 10.16.56.0 netmask 255.255.248.0 {
                option routers 10.16.56.1;
                option subnet-mask 255.255.248.0; } }
```

dhcpd.conf(Private)

```
shared-network hogekura{
       subnet A.B.0.0 netmask 255.255.255.0 {
               option routers A.B.0.1;
               option subnet-mask 255.255.255.0; }
       subnet 10.16.56.0 netmask 255.255.248.0 {
                option routers 10.16.56.1;
                option subnet-mask 255.255.248.0;
                pool {  allow members of  “hoge-master”;
                        allow members of  “hoge-client”;
                        range 10.16.56.2 10.16.63.254; } } }
```

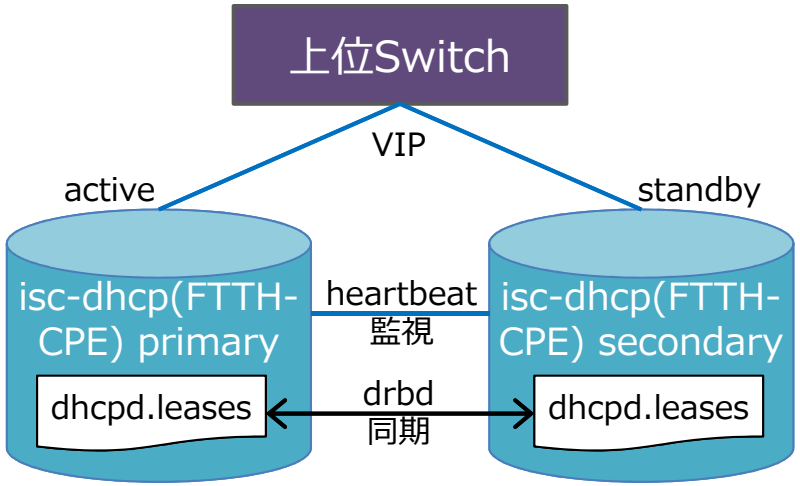
運用事例4 DHCPサーバの冗長

- isc-dhcp

標準でfailover機能がありますが、昔はあまり安定していないという話でしたが、今はわかりません、商用では使ったことがありません。

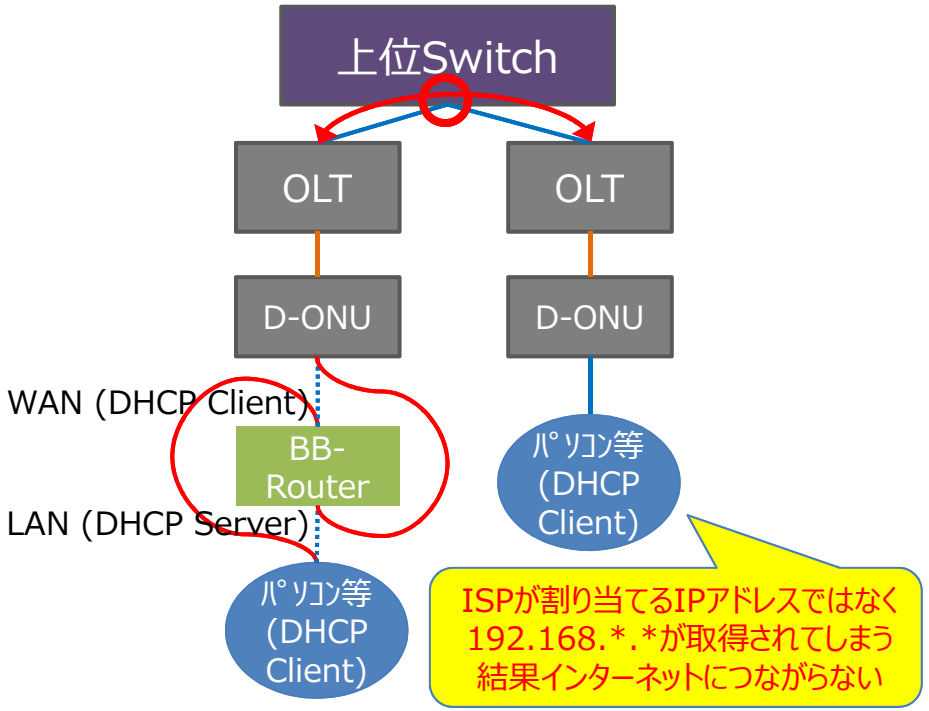
- drbd&heartbeat

OSSによる実装、dhcpd.leasesをdrbdで同期、heartbeatでプロセスを死活監視。



トラブル事例1 DHCPの逆流

- CPEのブロードバンドルータでWANとLANを逆にしてしまう場合がある。



回避策は以下など。

- 上位SwitchのDownLinkPortでswitchport protected
- PrivateVLAN
- ACLでudp67-68をin/outで適宜設定

トラブル事例2 isc-dhcp再起動が長い

- プライベートIPアドレスだからといって、/16のDHCP subnetをどんどん増やしていったら(/16*15くらい≒983k)、dhcpdの再起動に時間がかかるように(2-3分とかかかるのでもうドキドキ・・・)。
- どうもisc-dhcpはプロセスの起動時にIPアドレスの重複を防ぐために全てのPOOL rangeのIPアドレスにPINGで死活監視をしていたようだ。
- POOL rangeを実利用に合わせて縮退して対応。

```
dhcpd.conf
shared-network hogeshima {
        subnet 10.168.0.0 netmask 255.255.0.0 {
               option routers 10.168.0.1;
               option subnet-mask 255.255.0.0;
                pool {
                       range 10.168.192.2 10.168.255.255;
                } } }
...
```

トラブル事例3 いろいろなDHCP Client

- 短周期でDHCP Requestを投げてくるCPEあり。
- いろいろなDHCP Clientがあるので防げない(仕様?ユーザ設定?)。
- ISPがCPEルータまでmanageしてしまえば解決。コストもあるが、理想はIPv6/multicast対応WiFi内蔵ONUを全戸配布。合わせてBBF-TR069で遠隔サポート等もできると良い。

4~11秒間隔でDHCP Requestが届く

```
dhcpd.log | grep [CPE-MAC]
May 29 21:32:56 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 29 21:32:56 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
May 29 21:33:01 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 29 21:33:01 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
May 29 21:33:07 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 29 21:33:07 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
May 29 21:33:21 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 29 21:33:21 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
May 29 21:33:32 dhcp-ha1 dhcpd: DHCPREQUEST for [IPv4-address] from [CPE-MAC] via eth0
May 29 21:33:32 dhcp-ha1 dhcpd: DHCPACK on [IPv4-address] to [CPE-MAC] via eth0
```

IPv6対応

- いま進めてます。BGPでは広報済、末端までは未。
- 上位スイッチやOLTは IPv6 ready の状態(一部の古いOLTはOption18(Opt82のIPv6版)未対応)。
- 元々、終端装置(CMやD-ONU)でMACアドレス学習上限値を1とかにして、CPE接続台数制限をしている。
- OLTではLLID(Logical Link ID≠VLAN)毎に論理的にCPE接続台数制限できる製品もある。
- AndroidがDHCPv6未対応…。
- DHCPv6-PD(Prefix Delegation)方式ならMACアドレス学習上限数は増やさなくて良いので現実的な対応(ただしPD対応ルータが市場にはまだ少ない)。