



5G Innovations

IPv6トラブルシューティング

mini

---

株式会社ブロードバンドタワー  
國武 功一

## ■ IPv6 Troubleshooting!

- ◆ ipv6 nd ra suppress は、やめよう！
- ◆ おまけ - で、allなしコマンドの使い所は？ -
- ◆ おまけ - アドレス採番について
- ◆ おまけ - PMTUD Blackholeにまつわる誤解、もしくは懺悔
- ◆ おまけ - リンクローカルアドレスが付かない

あなたが書きたい設定は

```
ipv6 nd ra suppress
```

ではなく

```
ipv6 nd ra suppress all
```

ではないですか？

```
ipv6 nd ra suppress all
```

だよな？

```
ipv6 nd ra suppress all
```

というか、こう書こうぜ！

ご清聴ありがとうございました！

以上で、今日話したいことは、  
ほぼ語り尽くしました！

以下蛇足。

“all”をつけなくとも、定期的なRA送出は抑制されるので、一見、RA抑制ができたように思えます。

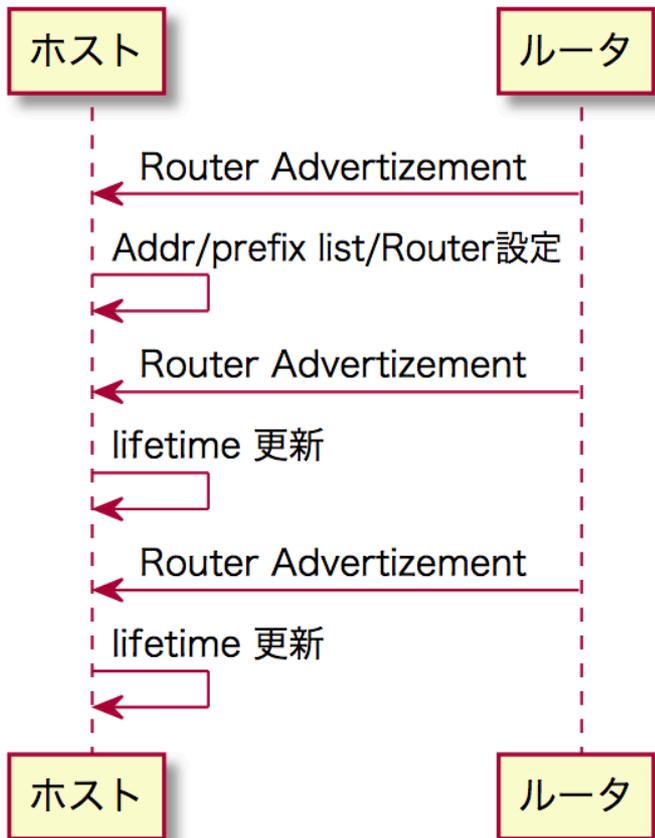
しかし、この状態でも、RS(Router Solicitation)パケットを受け取ると、RAを返してしまいます(\*1)

よくある？のは、Auto-configuration が On なノードがリンクを上げたタイミングで送出する RSパケットを受け取ってしまうケース。

<https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipv6/command/ipv6-cr-book/ipv6-i3.html?dtid=osscdc000283#wp2583862361>

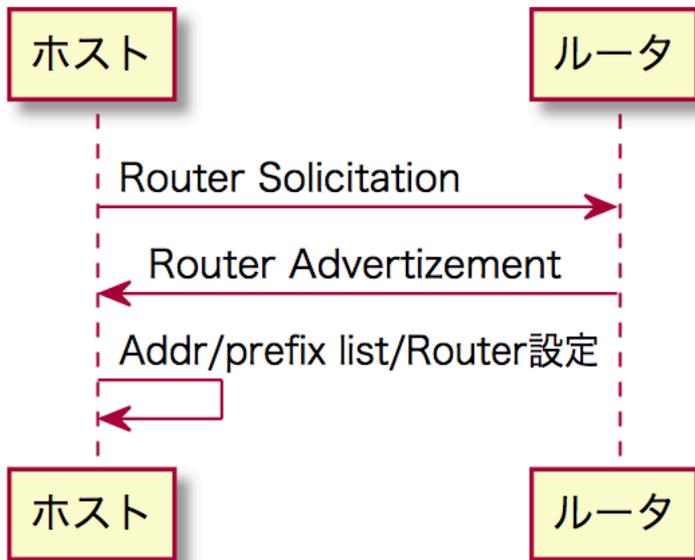
いつも良質なマニュアルをありがとう！~~けど、この箇所は分かりづらい~~

# 定期的に送出される RA (これは止まる)



Packet	Router Advertisement
Src mac	00:00:5e:00:53:ff
Dest mac	33:33:00:00:00:01
Src IPv6	fe80::200:5eff:fe00:53ff
Dest IPv6	ff02::1
Hoplimit	255
ICMPv6 Type	134
Prefix info	2001:db8::

# RSへの応答としてのRA (抑制にはALLが必要)



Packet	Router Solicitation
Src mac	00:00:5e:00:53:01
Dest mac	33:33:00:00:00:02 (IPv6mcast_02)
Src IPv6	fe80::200:5eff:fe00:5301
Dest IPv6	ff02::2
Hoplimit	255
ICMPv6 Type	133

Packet	Router Advertisement
Src mac	00:00:5e:00:53:ff
Dest mac	00:00:5e:00:53:01
Src IPv6	fe80::200:5eff:fe00:53ff
Dest IPv6	fe80::200:5eff:fe00:5301
Hoplimit	255
ICMPv6 Type	134
Prefix info	2001:db8::

従来の実装では、RSは基本的にnodeが I/F を up した時にしか流れない(\*1)ので、このRAでついたアドレスを利用すると、valid lifetimeが過ぎたあとに通信ができなくなる危険性がある（なにも設定変更していないと、30日後(\*2))

(\*1) RFC7559があるが、これは初期化時のRAの packet loss を想定した仕様

(\*2) default のパラメータはRFC4861に従っている実装が多い。

## ■ とある Linux の場合

1. RAを受け取る状態にしておく
2. RSに対するRAを受け取る。
3. eui64 のアドレスがつく (2001:db8::a00:27ff:feb4:76)
4. 2001:db8::/64 な経路(Prefix listに追加)が作られる。
5. static で固定のアドレスを /64 で割り当てる。  
(2001:db8::cafe/64)
6. default gatewayとして 2001:db8::/64 のアドレスを使って、nexthopを指定

## ■ とある Linux の場合(続き)

7. その後 RA が流れない
8. valid lifetime経過後、eui64なアドレスが消える
9. static で割り当てたアドレスが残っていてもprefix list から 2001:db8::/64 な経路が消える。
10. 結果、疎通性が無くなる(kernel の version によっては、RA受ける前にstatic経路を書いておけば、expires 付きの経路の登録を防げるケースも)

**サーバ屋さんとネットワーク屋さんとの  
華麗な共演 (設定ミス) による障害の発現**

RFC4861 6.3.4

*Prefix List - A list of the prefixes that define a set of addresses that are on-link. **Prefix List entries are created from information received in Router Advertisements.** Each entry has an associated invalidation timer value (extracted from the advertisement) used to expire prefixes when they become invalid. A special "infinity" timer value specifies that a prefix remains valid forever, unless a new (finite) value is received in a subsequent advertisement.*

某氏にお聞きしたところ、RA以外で生成された prefix は prefix listのエントリではない解釈をしている実装があれば影響を受けない、かもとのこと。

RA止めたいなら、細かいことは  
気にせず

```
ipv6 nd ra suppress all
```

と書いとけばOK！

おまけ - で、allなしコマンドの使い所は？ -

---

## ■ ISATAP(もう使っていないよね?)

- ◆ ISATAPでは、tunnel を通じて RA を流します。ルータでは tunnel を通じて multicast パケットを流すのが難しいので、RS にだけ答えるモードが求められていたそうです。

## ■ Wi-FiのGateway

- ◆ 無線ではなるべく、flooding するようなパケットを抑制したいので、RS にだけ答えるモードを使うという選択をするケースがあるようです (これが本当に運用に耐えるかどうかは、ユースケースに依存するでしょうけど……)

おまけ - アドレス採番について -

---

## ■ 2001:db8::dead:beaf

◆ このアドレス宛の NS は ff02::1:ffad:beaf がdestination

## ■ 2001:db8:aaad:beaf

◆ このアドレス宛の NS も ff02::1:ffad:beaf がdestination

■ たまたま同じmulticast groupにjoinしているルータが自分宛でないにも関わらず破棄せず、また同じセグメントに送り出しちゃうやつがいた(もちろんバグ)

## ■ IPv6で多様されるマルチキャストパケット

- ◆ 基本的には負荷低減。ホスト側で関係ないパケットを簡単に捨てられるようにしている。
- ◆ 同じマルチキャストグループに属していることがある、ということ意識していれば、トラブルシューティングの際に役に立つ……かも？

```
$ ip address show dev eth1
3:          eth1
    link 01:00:5e:00:00:01
    link 33:33:00:00:00:01
    link 33:33:ff:d3:66:67
    inet 224.0.0.1
    inet6 ff02::1:ffd3:6667
    inet6 ff02::1
    inet6 ff01::1
```

# おまけ PMTUD Blackholeにまつわる誤解

---

## IPv6 PMTU Discovery Blackholeの盲点

### 概要

IPv6を長年運用しているはずの組織でも、IPv6 PMTU Discovery Blackhole(なぜか? その盲点と再発防止策について、運用面から語ります。

IPv6 では中継ノードでフラグメントしない(始点ノードが実施)  
– IPv4 ではルータ等の中継ノードがフラグメントを実施

- IPv4と違って、IPv6では途中ルータがパケットをフラグメントすることは禁止されており、PMTU Discoveryが動作する  
マレトスと西に女期待してはため

<https://www.janog.gr.jp/meeting/janog34/doc/janog34-6pmtu-kunitake-1.pdf>

たとえばLinuxサーバでは DFビット立ててパケット投げてるので、IPv4でもルータではフラグメントしません

`ip_no_pmtu_disc` – INTEGER

Disable Path MTU Discovery. If enabled in mode 1 and a fragmentation-required ICMP is received, the PMTU to this destination will be set to `min_pmtu` (see below). You will need to raise `min_pmtu` to the smallest interface MTU on your system manually if you want to avoid locally generated fragments.

default で 0 なので、DFビット立てて送ります。1なら、PMTUDをdisableにします。つまりフラグメント期待してDFビット立てません。

<https://www.kernel.org/doc/Documentation/networking/ip-sysctl.txt>

- IPv4/IPv6もPMTUDが動作しないとネットワークは動きません。
- え？動いてるけど……？
  - ◆ TCP MSS Clamping で検索してみよう～

おまけ link-local addressが見つからない！

---

- 最近、Linux系ディストリビューションでRFC7217, RFC8064を意識した実装になってきている。
  - ◆ アドレスの下位64bitが Modified EUI-64ベースではない

## ■ 前提：RHEL7/CentOS7系のお話

- ◆ NetworkManagerがON
- ◆ IPV6INIT=yes
- ◆ IPV6\_ADDR\_GEN\_MODE=stable-privacyになっている。
- ◆ VM Cloneしていて、かつUUIDが同じ

```
IPV6INIT=yes
IPV6_ADDR_GEN_MODE=stable-privacy
UUID=9c92fad9-6ecb-3e6c-eb4d-8a47c6f50c04
TYPE=Ethernet
PROXY_METHOD=none
BROWSER_ONLY=no
```

## ■ 1台目

- ◆ stable-privacy な link-localアドレスがつく

## ■ 2台目

- ◆ 1台目と同じ stable-privacy な link-localアドレスが、DAD failed
- ◆ 新しい stable-privacyなアドレスを生成……link-localアドレスがつく

## ■ 3台目

- ◆ 1台目と同じ stable-privacy な link-localアドレスが、DAD failed
- ◆ 2台目と同じ stable-privacy な link-localアドレスが、DAD failed
- ◆ 新しい stable-privacyなアドレスを生成……link-localアドレスがつく

## ■ 4台目

- ◆ 1台目と同じ stable-privacy な link-localアドレスが、DAD failed
- ◆ 2台目と同じ stable-privacy な link-localアドレスが、DAD failed
- ◆ 3台目と同じ stable-privacy な link-localアドレスが、DAD failed
  - stable-privacyなアドレス生成を諦める
  - NSが投げられなくて詰む



**5G Innovations**