JANOG49 Meeting in Kagoshima

IXにおけるIPv6パケットフィルタ再考

日本インターネットエクスチェンジ
技術部　國武功一

2022-01-26

お断り

- 個人の見解を示した資料であり、組織を代表したものではなく云々……
- 経路フィルタについては扱いません
- IPv4パケットフィルタについても扱いません
- MACアドレスは文章用ドキュメントを利用していません。わかりやすさを優先させました。

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Agenda

- IXセグメントで見かけた不思議な挙動
- Neighbor Discoveryを正しく知ろう！
- JANOG Comment(JC1006)を振り返ろう！
- 勝手にJC1006まとめ
- JC1006で防げないへんなパケット例その1
- JC1006で防げないへんなパケット例その2
- さて、どうする？

Copyright © 2022 Japan Internet Exchange Co., Ltd.

フィルタしちゃダメなパケットまで
フィルタしてませんか？

IXセグメントで見かけた不思議な挙動

- Neighbor Discovery が失敗する
  - 時々成功
    - どうも Path MTU Discovery が定常的に走っていて、Rate limitに引っかかってる？
      - 参考：頑張れフォールバック – JANOG p.20 (IIJ 松崎さん, JANOG26)
  - 監視装置からの Neighbor Discovery は成功。対向ルータからのNeighbor Discoveryは失敗。またはその逆？
    - ？？？

Neighbor Discoveryで実際にどんなパケットが送受信されているかの理解がされず、誤ったフィルタが適用されているのでは？

Neighbor Discoveryを正しく知ろう！

- Neighbor Discoveryとは？
  - ざっくり言えば、IPv4でいうところの、ARPみたいなもの
  - ICMPv6 を利用する(Neighbor SolicitationとNeighbor Advertisement)
- どんな時に投げる？
  - L2アドレスの解決(ARP相当だよ）
  - アドレス重複の確認(DAD)
  - 近隣到達不能検出(NUD) の確認
  - Bondingでよく使われるGARP相当のパケット(切り替わった際に、新しいポートにMACアドレスを再学習させる）

Copyright © 2022 Japan Internet Exchange Co., Ltd.

RFC4861 Neighbor Discovery for IP version 6

4.3. Neighbor Solicitation Message Format

Nodes send Neighbor Solicitations to request the link-layer address of a target node while also providing their own link-layer address to the target. Neighbor Solicitations are multicast when the node needs to resolve an address and unicast when the node seeks to verify the reachability of a neighbor.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Code      |          Checksum             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Reserved                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                       Target Address                          +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Options ...
+-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
Either an address assigned to the interface from which this message is sent or (if Duplicate Address Detection is in progress [ADDRCONF]) the unspecified address.

Destination Address
Either the solicited-node multicast address corresponding to the target address, or the target address.

Hop Limit   255

以下略

4.4. Neighbor Advertisement Message Format

A node sends Neighbor Advertisements in response to Neighbor Solicitations and sends unsolicited Neighbor Advertisements in order to (unreliably) propagate new information quickly.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Code      |          Checksum             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|S|O|                     Reserved                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                       Target Address                          +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Options ...
+-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
An address assigned to the interface from which the advertisement is sent.

Destination Address
For solicited advertisements, the Source Address of an invoking Neighbor Solicitation or, if the solicitation's Source Address is the unspecified address, the all-nodes multicast address.

For unsolicited advertisements typically the all-nodes multicast address.

Hop Limit   255

以下略

Copyright © 2022 Japan Internet Exchange Co., Ltd.

IPv6 src/dst アドレス

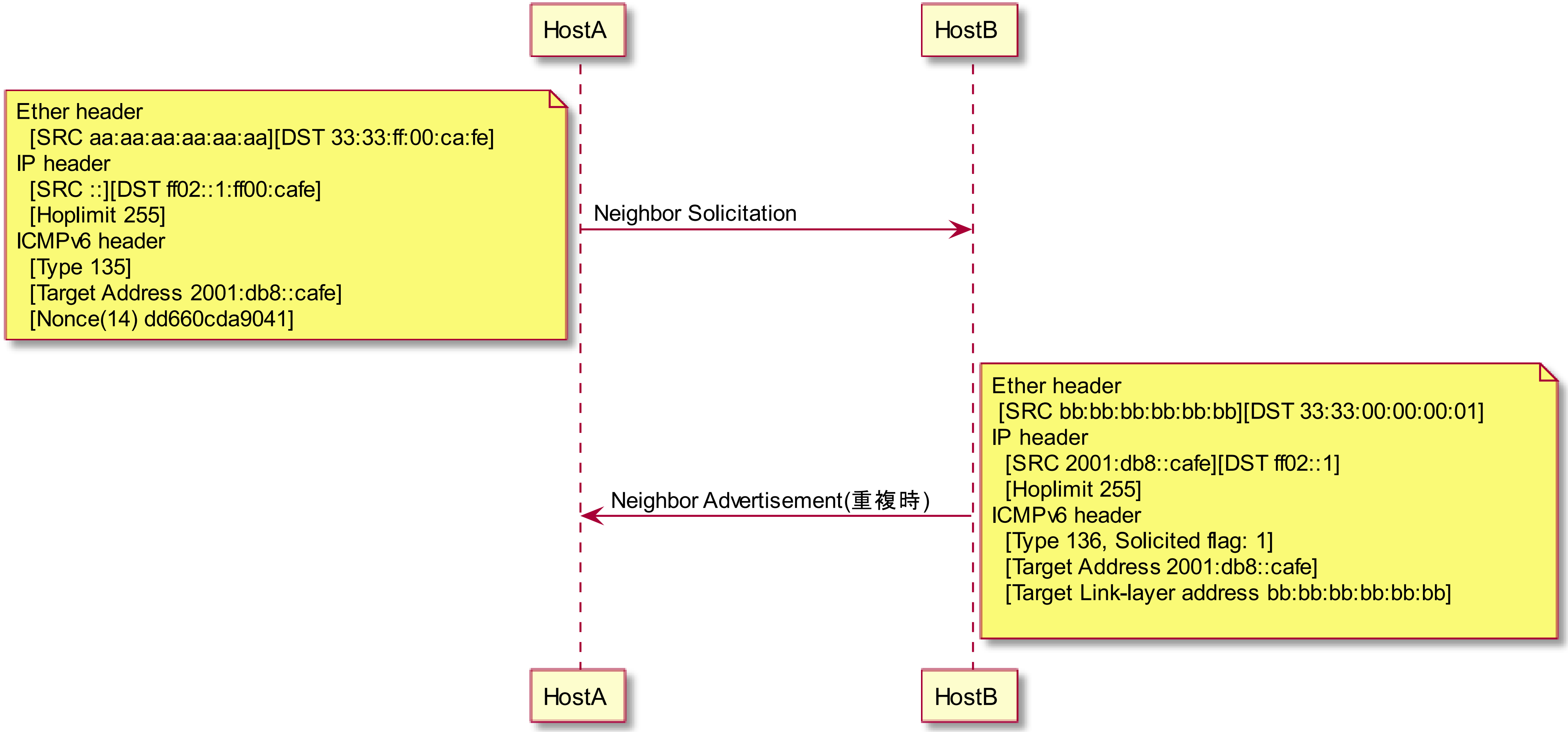
- Neighbor Solicitation
  - IPv6 Source Address : Interfaceにアサインされた（global, link-local, DADなら未指定アドレス(::))アドレス。
  - IPv6 Destination Address : 解決したいターゲットアドレスの要請ノードマルチキャストアドレス(Solicited-Node Multicast address)かターゲットアドレスそのもの
- Neighbor Advertisement
  - IPv6 Source Address : 送出するInterfaceのIPアドレス(globalやらlink-localアドレスやら)
  - IPv6 Destination Address: NSの送信元アドレスや DADへの応答は全ノードマルチキャスト宛。

Copyright © 2022 Japan Internet Exchange Co., Ltd.

具体的には？

Neighbor Solicitation その１

- DAD(Duplicate Address Detection)



HostA

HostB

Ether header
[SRC aa:aa:aa:aa:aa:aa][DST 33:33:ff:00:ca:fe]
IP header
[SRC ::][DST ff02::1:ff00:cafe]
[Hoplimit 255]
ICMPv6 header
[Type 135]
[Target Address 2001:db8::cafe]
[Nonce(14) dd660cda9041]

Neighbor Solicitation

Ether header
[SRC bb:bb:bb:bb:bb:bb][DST 33:33:00:00:00:01]
IP header
[SRC 2001:db8::cafe][DST ff02::1]
[Hoplimit 255]
ICMPv6 header
[Type 136, Solicited flag: 1]
[Target Address 2001:db8::cafe]
[Target Link-layer address bb:bb:bb:bb:bb:bb]

Neighbor Advertisement(重複時)

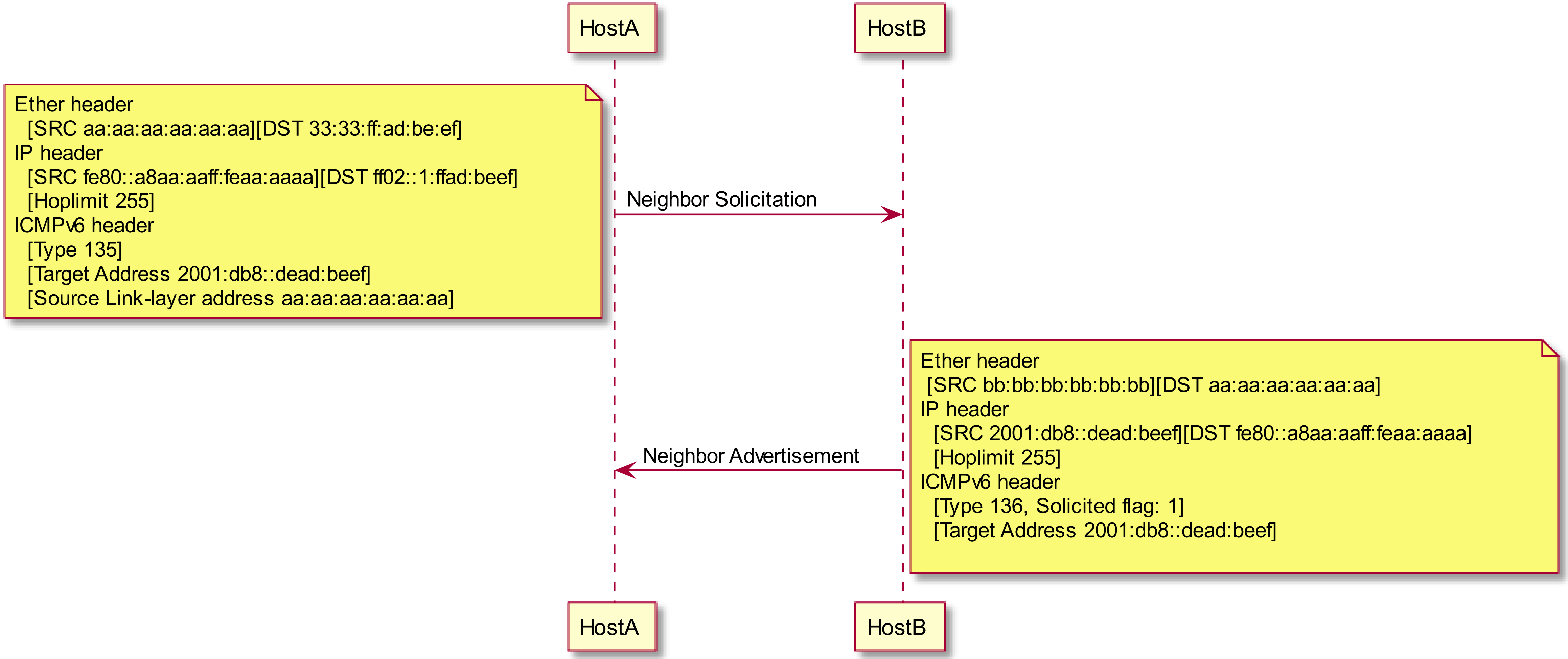
HostA

HostB

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Neighbor Solicitation その２

- L2アドレス解決(link-local -> solicited node multicast)



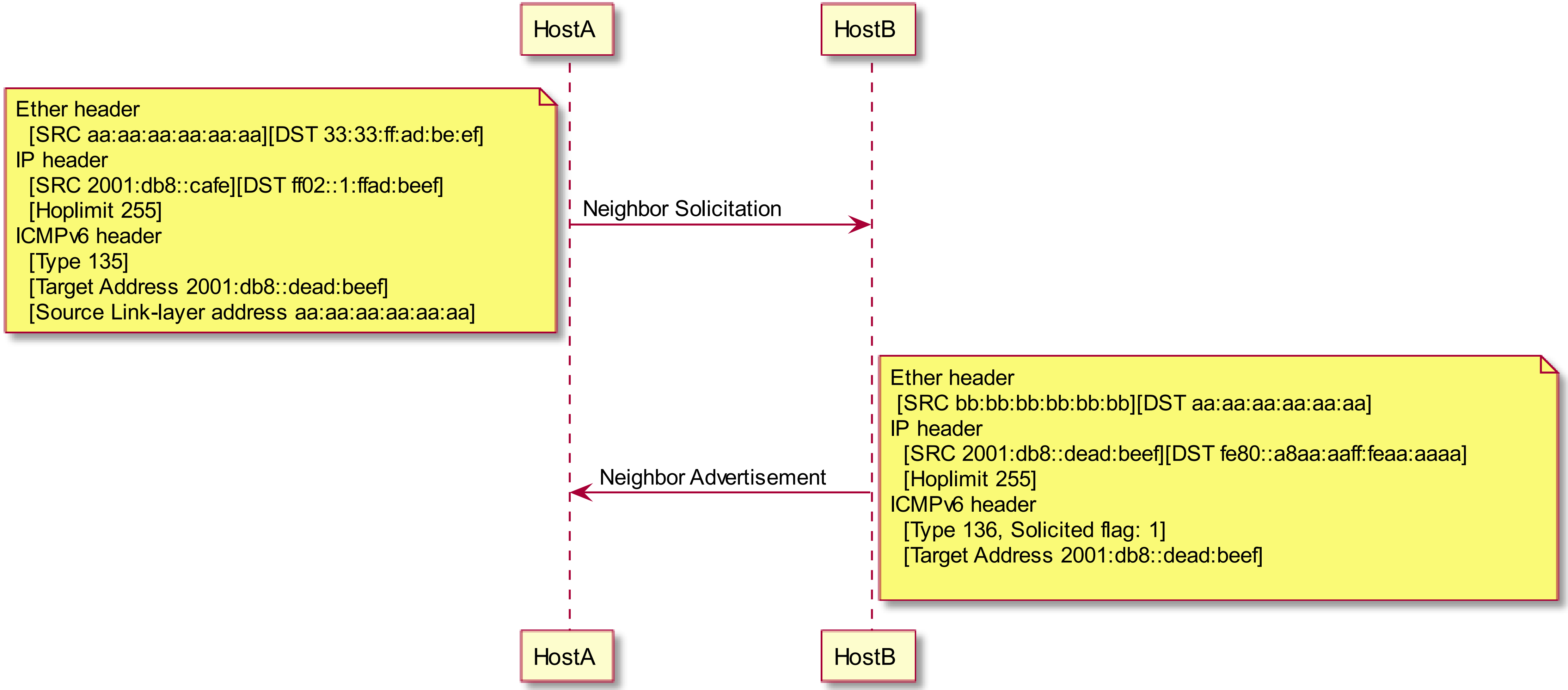
Ether header
  [SRC aa:aa:aa:aa:aa:aa][DST 33:33:ff:ad:be:ef]
IP header
  [SRC fe80::a8aa:aaff:feaa:aaaa][DST ff02::1:ffad:beef]
  [Hoplimit 255]
ICMPv6 header
  [Type 135]
  [Target Address 2001:db8::dead:beef]
  [Source Link-layer address aa:aa:aa:aa:aa:aa]

Ether header
  [SRC bb:bb:bb:bb:bb:bb][DST aa:aa:aa:aa:aa:aa]
IP header
  [SRC 2001:db8::dead:beef][DST fe80::a8aa:aaff:feaa:aaaa]
  [Hoplimit 255]
ICMPv6 header
  [Type 136, Solicited flag: 1]
  [Target Address 2001:db8::dead:beef]

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Neighbor Solicitation その３

- L2アドレス解決(global -> solicited node multicast)



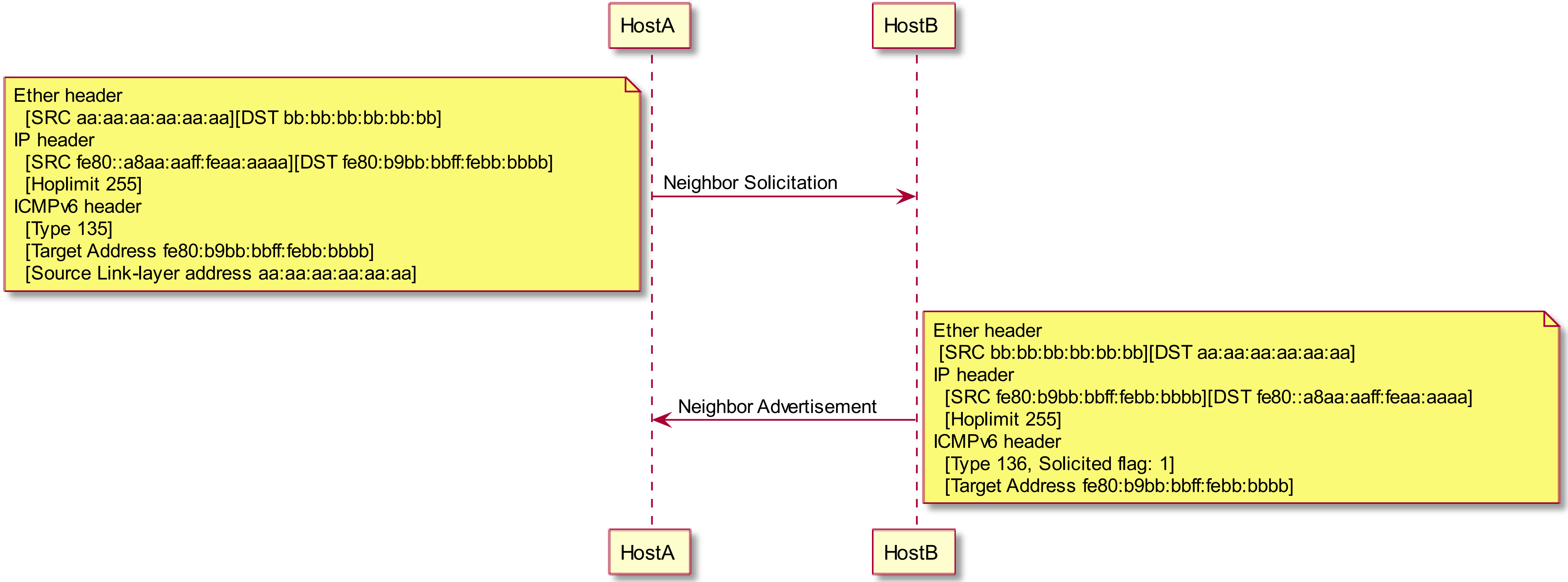
Ether header
[SRC aa:aa:aa:aa:aa:aa][DST 33:33:ff:ad:be:ef]
IP header
[SRC 2001:db8::cafe][DST ff02::1:ffad:beef]
[Hoplimit 255]
ICMPv6 header
[Type 135]
[Target Address 2001:db8::dead:beef]
[Source Link-layer address aa:aa:aa:aa:aa:aa]

Ether header
[SRC bb:bb:bb:bb:bb:bb][DST aa:aa:aa:aa:aa:aa]
IP header
[SRC 2001:db8::dead:beef][DST fe80::a8aa:aaff:feaa:aaaa]
[Hoplimit 255]
ICMPv6 header
[Type 136, Solicited flag: 1]
[Target Address 2001:db8::dead:beef]

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Neighbor Solicitation その３

- NUD(Neighbor Unreachability Detection)



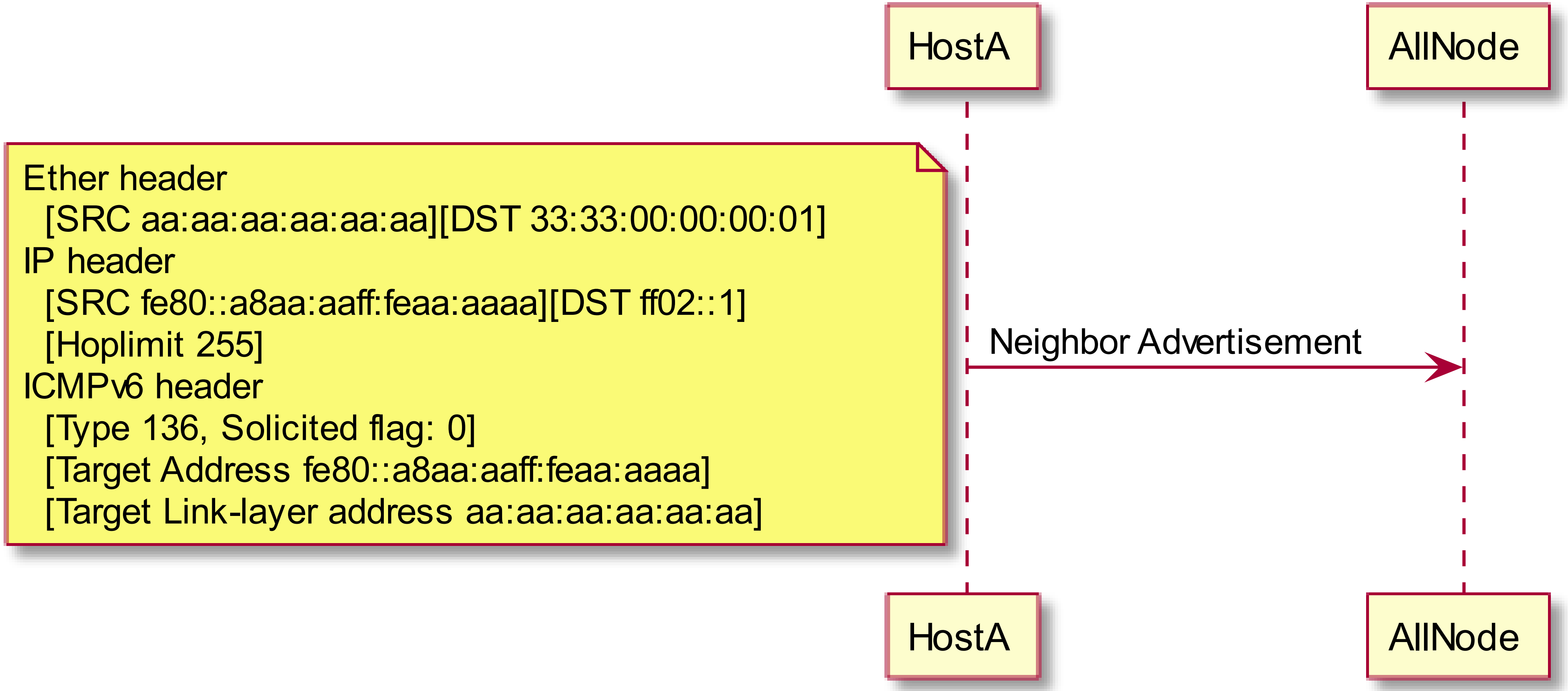
Ether header
  [SRC aa:aa:aa:aa:aa:aa][DST bb:bb:bb:bb:bb:bb]
IP header
  [SRC fe80::a8aa:aaff:feaa:aaaa][DST fe80:b9bb:bbff:febb:bbbb]
  [Hoplimit 255]
ICMPv6 header
  [Type 135]
  [Target Address fe80:b9bb:bbff:febb:bbbb]
  [Source Link-layer address aa:aa:aa:aa:aa:aa]

Ether header
 [SRC bb:bb:bb:bb:bb:bb][DST aa:aa:aa:aa:aa:aa]
IP header
  [SRC fe80:b9bb:bbff:febb:bbbb][DST fe80::a8aa:aaff:feaa:aaaa]
  [Hoplimit 255]
ICMPv6 header
  [Type 136, Solicited flag: 1]
  [Target Address fe80:b9bb:bbff:febb:bbbb]

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Neighbor Advertisement 単体

- GARPみたいなやつ



Ether header
  [SRC aa:aa:aa:aa:aa:aa][DST 33:33:00:00:00:01]
IP header
  [SRC fe80::a8aa:aaff:feaa:aaaa][DST ff02::1]
  [Hoplimit 255]
ICMPv6 header
  [Type 136, Solicited flag: 0]
  [Target Address fe80::a8aa:aaff:feaa:aaaa]
  [Target Link-layer address aa:aa:aa:aa:aa:aa]

Copyright © 2022 Japan Internet Exchange Co., Ltd.

The Core of
Internet Community

などなど……

StrictなICMPv6パケットのフィルタは十分複雑

- IPv4感覚のフィルタでは誤りにつながる。
  - IXセグメントのアドレスだけを許可、ではダメ。
  - All-Node Multicast, Solicited Node Multicast や link-local address もsrc/dst アドレスとして必要に応じて許可する必要がある(src に限れば、未指定アドレスも許可する必要もある）
    - Destination アドレスが特に複雑

Copyright © 2022 Japan Internet Exchange Co., Ltd.

JANOG Comment 振り返ろう！

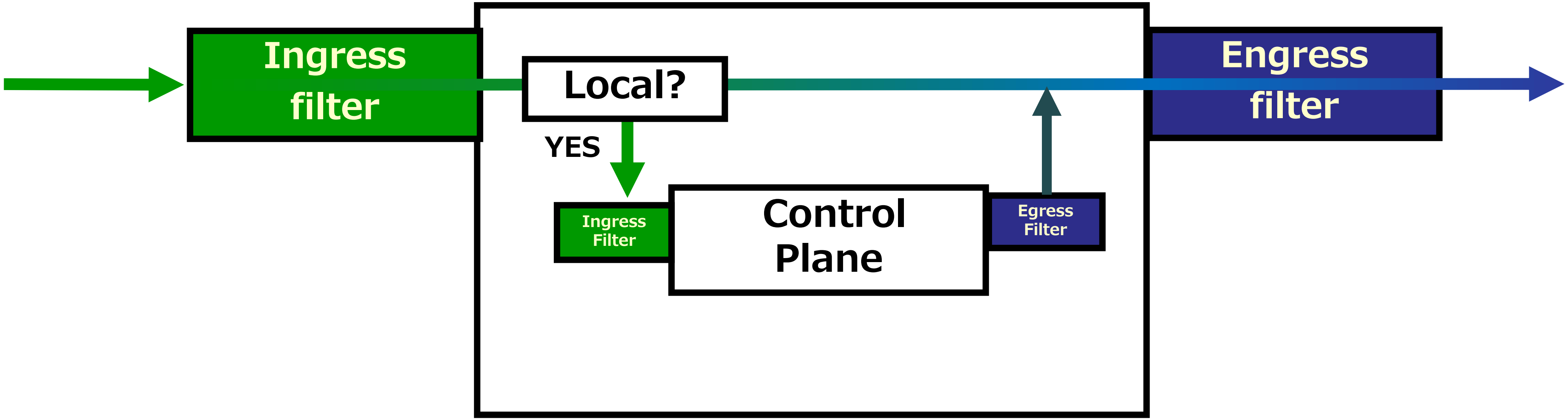
JC1006

JANOG Comment 1006 xSP のルータにおいて設定を推奨するフィルタの項目について（IPv6版）

- トランジット接続部分
- **ピア接続部分**
- 顧客接続部分
- ルータ自身へのアクセス部分

Interfaceに設定するPacket filter（トランジット接続部分・ピア接続部分・顧客接続部分）

Control Planeに設定するPacket filter（ルータ自身へのアクセス部分）



Copyright © 2022 Japan Internet Exchange Co., Ltd.

I/FへのIngressのパケットフィルタ（JC1006)

- ICMPv6は基本 Accept
  - ND(Neighbor Discovery)用, Path MTU Discovery, Destination unreachable
- あり得ないSrc IP Addressが付いたパケットはフィルタ
- 顧客接続については、本プログラムでは省略

		Ingress											
		ICMPv6	下記 Special-Use PrefixがSource IPなパケット										自ASが持っているアドレスが source のパケット
			::/96	::1/128	::ffff:0:0/96	100::/64	2001:2::/48	2001:db8::/32	fc00::/7	fec0::/10	ff00::/8	::/128	
トランジット	推奨	permit all	drop									Allow (ICMPv6 DADで利用)	drop(UDLR/非対称ルーティングの場合に注意)
	余裕があれば	strict											
ピア	推奨	permit all											
	余裕があれば	strict											
顧客接続	推奨	permt all											drop(トランジット顧客の場合)
	余裕があれば	strict											顧客の側でもってるprefix のみをaccept 自 AS と接続をしている IX セグメントのアドレスがDestination アドレスとなっている BGP (179/TCP) パケットをdrop

I/FへのEgressのパケットフィルタ（JC1006)

- 基本推奨設定なし

		リソースに余裕があれば下記 Special-Use PrefixがSource IPなパケット									
		::/96	::1/128	::ffff:0:0/96	100::/64	2001:2::/48	2001:db8::/32	fc00::/7	fec0::/10	ff00::/8	::/128
トランジット ピア 顧客接続	推奨	N/A									
	余裕があれば	drop									allow(ICMPv6) DADで利用
ピア	推奨	N/A									
	余裕があれば	drop									allow(ICMPv6) DADで利用
顧客接続	推奨	N/A									
	余裕があれば	drop									allow(ICMPv6) DADで利用

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Control PlaneへのIngress フィルタ(JC1006)

- 推奨
  - telnet/ssh/snmp/ftp/tftp/ntpなど、必要なサービスに対し、限定した source IPアドレスからのパケットのみ accept
  - eBGP/iBGPの neighorアドレスが sourceのBGPパケットのみを accept
  - Source address が neighbor の link-localアドレスとなっているパケットを accept (NDの動作で必要）
- 余裕があれば
  - ICMPv6 TYPEを制限して accept (I/FへのIngressフィルタと重複）

Copyright © 2022 Japan Internet Exchange Co., Ltd.

Control PlaneへのEgress フィルタ(JC1006)

- 推奨
  - 特になし
- 余裕があれば
  - 特になし

Copyright © 2022 Japan Internet Exchange Co., Ltd.

勝手にJC1006まとめ

- Interfaceへのパケットフィルタ
  - ICMPv6は基本全部通す。
  - 変な送信元IPv6アドレスのパケットを受け取らない
- Control Planeへのパケットフィルタ
  - ICMPv6は基本受け取る。
  - 利用サービスのパケットは受信できるようにする。送信元、宛先が仮定できるならIPv6アドレスベースでもフィルタもする。
- 現状も十分通用するように見える。
  - フィルタ粒度もリーズナブルに思える。

Copyright © 2022 Japan Internet Exchange Co., Ltd.

JC1006で万事解決！

- と思っていた時期もありました。
- 不思議な挙動を見かけた

へんなパケット(Neighbor Solicitation)例 その１

- Destinationが同一の Solicited-Node multicast address(要請ノードマルチキャストアドレス)のケースで、実際は自分宛ではなかった時の、ある実装の挙動

RouterA 2001:db8::bead:beef

RouterB 2001:db8::dead:beef

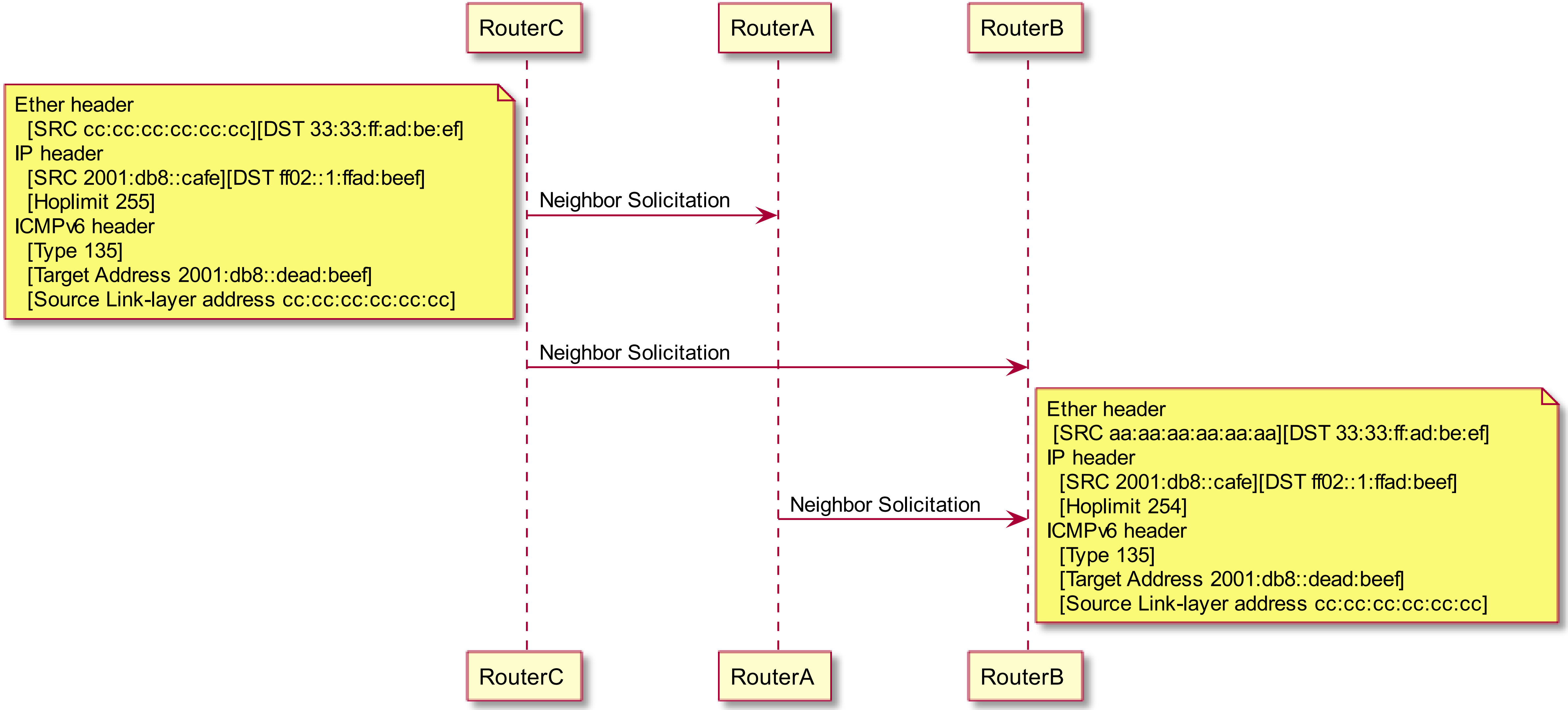
上記２アドレスともに

Solicited Node Multicast(ff02::1:ff00:0/104)

ff02::1:ffad:beef

Copyright © 2022 Japan Internet Exchange Co., Ltd.

RouterAは自分宛でないと判断し転送



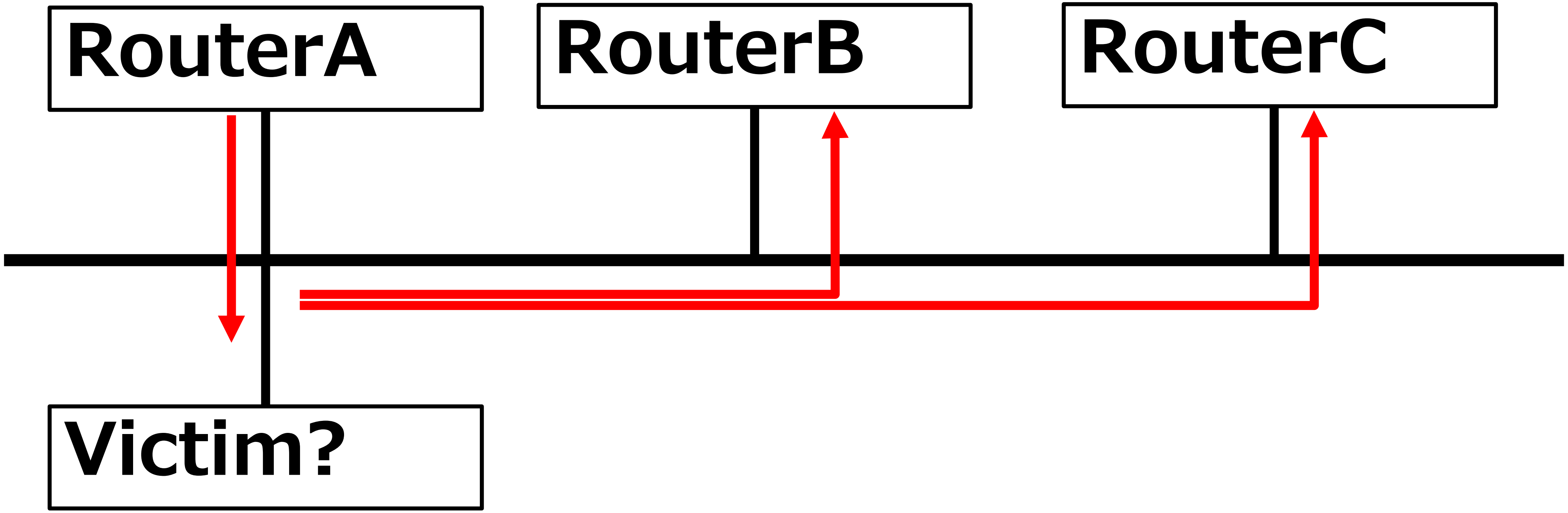
Ether header
  [SRC cc:cc:cc:cc:cc:cc][DST 33:33:ff:ad:be:ef]
IP header
  [SRC 2001:db8::cafe][DST ff02::1:ffad:beef]
  [Hoplimit 255]
ICMPv6 header
  [Type 135]
  [Target Address 2001:db8::dead:beef]
  [Source Link-layer address cc:cc:cc:cc:cc:cc]

Ether header
  [SRC aa:aa:aa:aa:aa:aa][DST 33:33:ff:ad:be:ef]
IP header
  [SRC 2001:db8::cafe][DST ff02::1:ffad:beef]
  [Hoplimit 254]
ICMPv6 header
  [Type 135]
  [Target Address 2001:db8::dead:beef]
  [Source Link-layer address cc:cc:cc:cc:cc:cc]

Copyright © 2022 Japan Internet Exchange Co., Ltd.

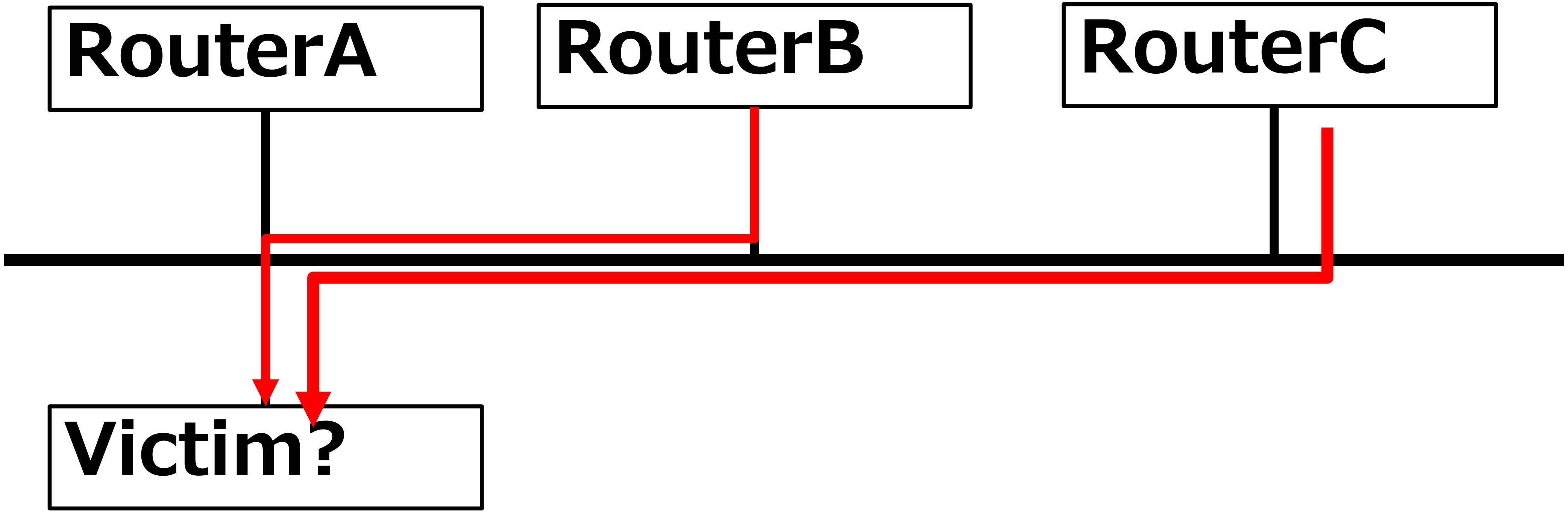
へんなパケット(Neighbor Solicitation)例 その２

- Destination が Victim 宛て。ただし、destination の MAC アドレスが、unknown unicast 宛てになっているので、全ポートにそのパケットが届く



へんなパケット(Neighbor Solicitation)例 その２

- パケットを受け取ったルータは、そのパケットが自分宛ではないので、ルータの本能？に従って転送。ちょっとした増幅攻撃に。



Copyright © 2022 Japan Internet Exchange Co., Ltd.

なぜ起きる？ルータの本能？

- 「自分宛じゃないパケットは経路表に従い転送する」
- IPv6の multicastってちょっと一部特殊かも
  - 受信処理の計算量を減らすために、作られている multicastアドレスがある(Solicited-Node multicast)
  - このアドレスは受け取っても自分宛でない可能性がある。で、本能？に従い転送
- だけどこれってバグだよね……
  - ルータベンダーさんはソフトウェア処理は重いと言われますがこれって、ASICで排除できたりしますかね……

Copyright © 2022 Japan Internet Exchange Co., Ltd.

どうすればいいの？

- 受け取っても実害はないので無視するか、hoplimit=255ではない NSパケットは Egress filterでdrop する（他人に迷惑をかけない）
- が、バグに類する挙動だと思われるので、フィルタでの対処は推奨したくない（どうせ受信側では drop されるし修正されるバグなら不要なフィルタとなる）

個人的な結論

- JC1006はやっぱり今でも通用する良いドキュメントだよ。
- Neighbor Discoveryはアドレスの組み合わせも多く、ルータベンダーにとっても複雑。へんな挙動を見かけたら、バグ報告しよう(ごめんなさい、今回紹介した変なNS２件、バグ報告してません。きっと既知の問題……だといいな)

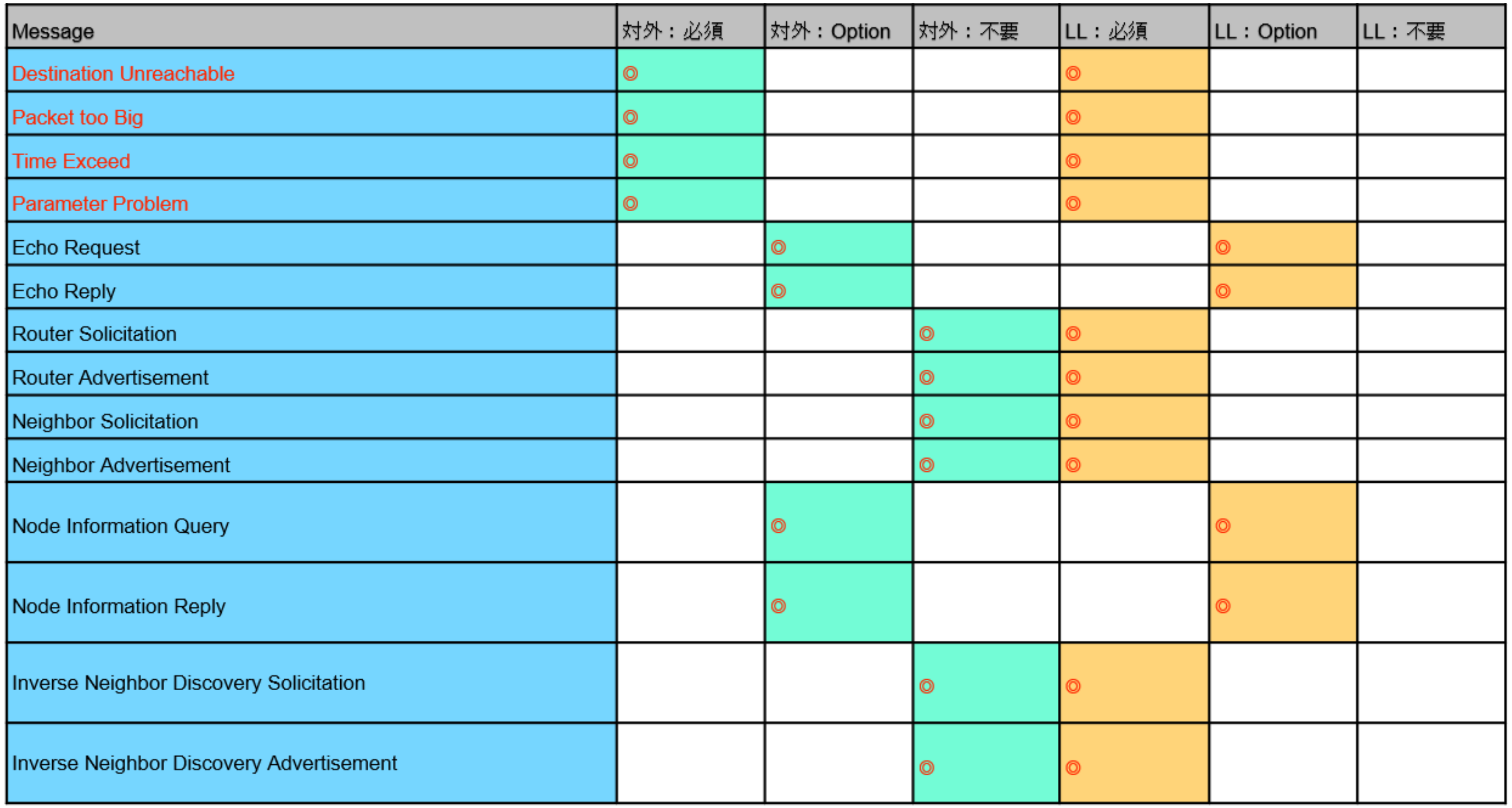
議論

- そもそもフィルタ掛けてますか？
- 細かくフィルタしたいですか？
  - 目の敵にされるICMPv6(ICMP)
    - 対象はコントロールプレーン向け？
    - 大昔あった、broadcastアドレスを使ったICMPの増幅アタックの記憶があとを引きずっている？
    - もし細かくフィルタをするのであれば、そのプロトコルへの深い理解が必要ですよね……
- JC1006に加えて、推奨したいフィルタはありますか？

Copyright © 2022 Japan Internet Exchange Co., Ltd.

参考：ICMPv6をTYPE毎にフィルタをお考えなら

ICMP6を「対外通信・内部通信」と「必須・オプション」に分けて考える



Message	対外：必須	対外：Option	対外：不要	LL：必須	LL：Option	LL：不要
Destination Unreachable	◎			◎		
Packet too Big	◎			◎		
Time Exceed	◎			◎		
Parameter Problem	◎			◎		
Echo Request		◎			◎	
Echo Reply		◎			◎	
Router Solicitation			◎	◎		
Router Advertisement			◎	◎		
Neighbor Solicitation			◎	◎		
Neighbor Advertisement			◎	◎		
Node Information Query		◎			◎	
Node Information Reply		◎			◎	
Inverse Neighbor Discovery Solicitation			◎	◎		
Inverse Neighbor Discovery Advertisement			◎	◎		

※JPNIC技術セミナー IPv6オペレータ育成プログラム 入門IPv6より

Copyright © 2022 Japan Internet Exchange Co., Ltd.

The Core of
Internet Community

Japan Internet Exchange Co., Ltd.

Copyright © 2022 Japan Internet Exchange Co., Ltd.