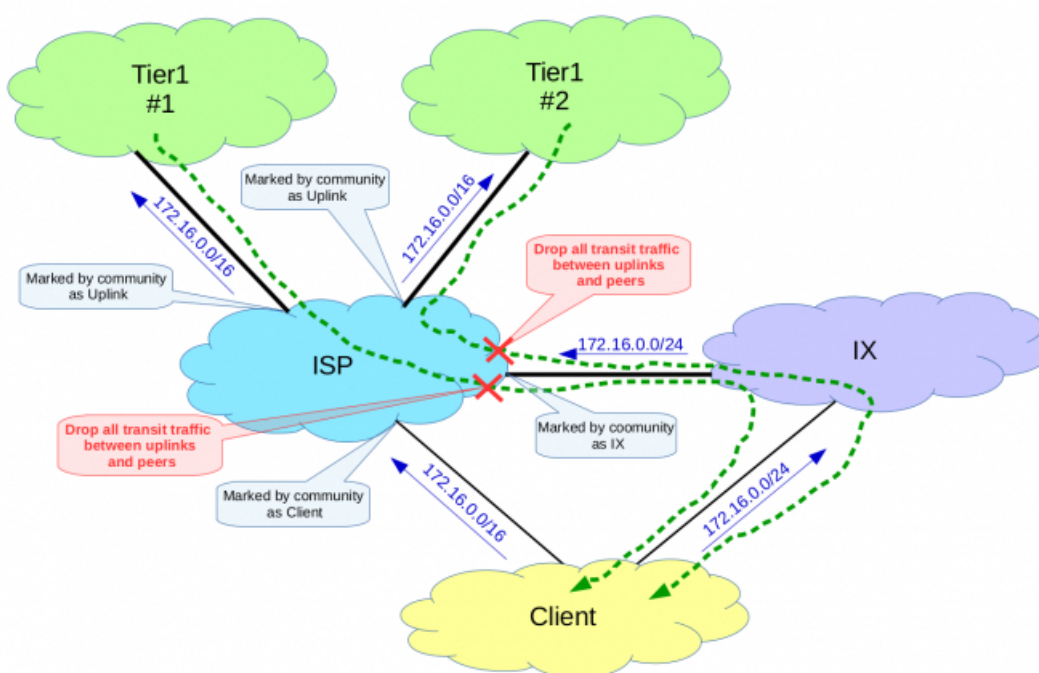


Игры с анонсами less specific и more specific, или из-за чего страдают клиенты

К написанию данной статьи меня сподвигли участвовавшие случаи обращения сторонних провайдеров с проблемами о недоступности части ресурсов отдельных наших клиентов.

Суть подобных проблем довольно проста, и заключается в желании клиентов получать более дешевый трафик с обменников с одной стороны, и нежелании ISPs осуществлять транзит трафика своих клиентов по схемам Uplink-->ISP-->IX-->Client, Uplink-->ISP-->Uplink-->Client, IX-->ISP-->IX-->Client с другой стороны.

Рассмотрим один из возможных вариантов на нижеприведенном рисунке.



Multihomed Client, имеющий подключения к ISP и IX, анонсирует менее специфичный префикс 172.16.0.0/16 своему ISP и более специфичный 172.16.0.0/24 на обменник. ISP принимает клиентский более специфичный маршрут 172.16.0.0/24 от IX, так как тоже является его участником. Принимаемый непосредственно от клиента префикс 172.16.0.0/16 ISP анонсирует своим аплинкам, не являющимися участниками пиринга на IX (как правило, это Tier1, или национальные Операторы, полностью отсутствующие на IX-ах, или присутствующие там со значительными ограничениями маршрутных политик). В результате подобной "кривизны" в анонсах трафик, предназначенный Client, из сетей Tier1 попадает в сеть ISP, откуда должен доставляться клиенту через IX по более специфичному маршруту 172.16.0.0/24, который будет всегда активным в сравнении с менее специфичным.

Кому хорошо в данной ситуации? Однозначно клиенту, получающему глобальную связность и трафик через более дешевый стык с IX. А кому плохо? Ответ очевиден, ISP, которому необходимо оплачивать транзит подобного клиентского трафика дважды, не имея какого-либо профита от самого клиента.

Какие существуют способы борьбы с подобными явлениями на сетях ISPs? Самым простым покажется фильтрация более специфичного префикса на сети ISP. На самом деле это далеко не так. Основной его проблемой является то, что подобные аномалии не являются статичными и их необходимо каким-то образом выявлять.

Другим, более распространенным способом, применяемым в сетях крупных ISPs, является выявление подобного трафика и его drop на форвардинге. Рассмотрим данный способ более подробно с примерами конфигурации для оборудования Juniper. Основной смысл его заключается, что префиксы, принимаемые от Uplinks и IXs, раскрашиваются BGP communities, на основании которых формируется, знакомый многим по предыдущим публикациям, destination-class. Сформированный destination-class и интерфейсы включения Uplinks и IXs являются основными критериями match для firewall фильтра с опцией then discard, который применяется на форвардинге.

Ниже приведены примеры конфигурации с короткими аннотациями:

1) Создаются community для раскраски принимаемых анонсов от аплинков и обменников:

```
set policy-options community mark-tier1_1 members 65001:30203
set policy-options community mark-tier1_2 members 65001:30204
set policy-options community mark-ix members 65001:20204
```

2) Созданные community включаются в policy, которое, в свою очередь, применяется на import в конфигурации протокола BGP для соответствующих групп аплинков и обменников:

```
set policy-options policy-statement mark-tier1_1 term mark-upstream then community add mark-tier1_1
set policy-options policy-statement mark-tier1_2 term mark-upstream then community add mark-tier1_2
set policy-options policy-statement mark-ix term mark-upstream then community add mark-ix
```

3) Создается community для определения ранее раскрашенных маршрутов:

```
set policy-options community type_peer-and-upstream members "^65001:[2,3]....$"
```

4) По определяющему community с помощью policy формируется destination-class:

```
set policy-options policy-statement set-class term peer-and-upstream from protocol bgp
set policy-options policy-statement set-class term peer-and-upstream from community type_peer-and-upstream
set policy-options policy-statement set-class term peer-and-upstream then destination-class peer-and-upstream
```

5) Policy с destination-class применяется на export для forwarding-table:

```
set routing-options forwarding-table export set-class
```

6) Создается firewall filter, основными критериями которого являются сформированный ранее destination-class и интерфейсы включения аплинков и обменников, а в качестве action данного фильтра discard и подсчет пакетов:

```
set firewall filter forward_firewall term drop-peers-to-other-peer from destination-class peer-and-upstream
set firewall filter forward_firewall term drop-peers-to-other-peer from interface ae7.0
set firewall filter forward_firewall term drop-peers-to-other-peer from interface ae6.0
set firewall filter forward_firewall term drop-peers-to-other-peer from interface ae5.0
set firewall filter forward_firewall term drop-peers-to-other-peer then count peers-to-other-peers-drop
set firewall filter forward_firewall term drop-peers-to-other-peer then discard
set firewall filter forward_firewall term any then accept
```

7) Созданный фильтр применяется на форвардинге:

```
set forwarding-options family inet filter output forward_firewall
```

Посмотреть срабатывание счетчиков данного фильтра можно командой:

```
> show firewall filter forward_firewall
Filter: forward_firewall
Counters:
Name                                     Bytes          Packets
peers-to-other-peers-drop               14704283695    182675892
{master}
```

В качестве заключения хочу выразить уверенность, что данная статья окажет помощь в понимании того, почему нельзя таким образом анонсировать более и менее специфичные маршруты, а также в том, что показания счетчиков дропнутых некорректно маршрутизируемых пакетов в сетях Операторов сократятся к нулю.