Measuring IXP Route Filtering

Doug Madory, Director of Internet Analysis
Garrett Allen, Senior Software Engineer

Global Peering Forum
Montreal, Canada April 2019

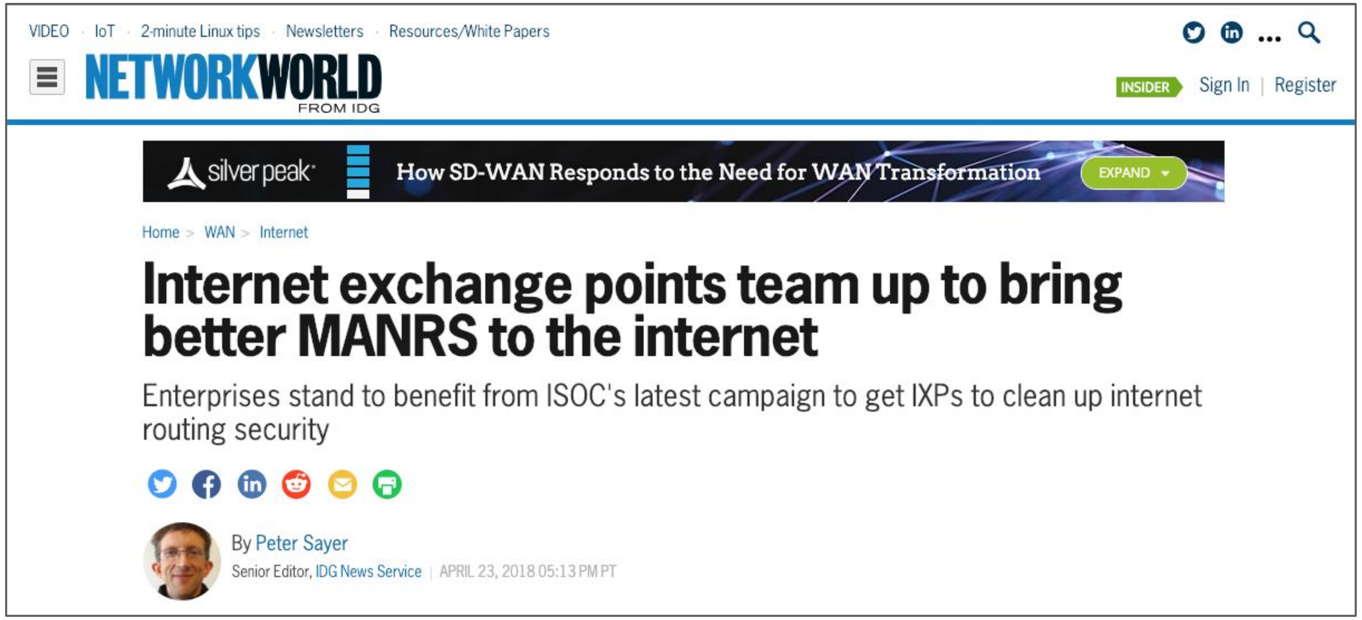
Unique Role of IXPs to Help Routing Security

- Improving global BGP routing hygiene is a notoriously difficult task.
  - Universal adoption of RPKI and other security mechanisms not likely
  - But incremental improvements are possible!

- IXP Route Servers offer an opportunity to filter inappropriate BGP messages
  - Based on best practice filters (RPKI/IRR invalids, bogons, etc.)
  - A value to their IXP members and for the good of the internet.

ORACLE

April 2018: Internet Society launches MANRS IXP Program


VIDEO · IoT · 2-minute Linux tips · Newsletters · Resources/White Papers

NETWORKWORLD FROM IDG

INSIDER Sign In | Register

silver peak How SD-WAN Responds to the Need for WAN Transformation EXPAND

Home > WAN > Internet

Internet exchange points team up to bring better MANRS to the internet

Enterprises stand to benefit from ISOC's latest campaign to get IXPs to clean up internet routing security

By Peter Sayer
Senior Editor, IDG News Service | APRIL 23, 2018 05:13 PM PT

Action 1. Prevent propagation of incorrect routing information. (Mandatory)

https://www.manrs.org/ixps/

MANRS IXP Program(me)

The IXP Programme Action Set

Action 1. Prevent propagation of incorrect routing information. (Mandatory)

The IXP implements filtering of route announcements at the Route Server based on routing information data (IRR and/or RPKI). Based on the outcome of the validation process, the invalid announcements are filtered in accordance with the IXP published policy.

IXPs using a Route Server to facilitate multilateral peerings should use it to validate received route announcements from a peer and subsequently filter them to other peers. Special purpose cases, such as research projects, are out of scope for this requirement.

Validation is usually done by checking BGP announcements against IRR data (by resolving the AS-SET object) or RPKI data (ROA objects or a validated cache). It is also common to check the announcements against “bogons” or “martians” (IP prefixes as defined in RFC1918, RFC5735, and RFC6598; ASNs in the AS-PATH as defined by RFC5398, RFC6793, RFC6996, RFC7300, RFC7607).

https://www.manrs.org/ixps/#actions

ORACLE

Can we measure this?

> Validation is usually done by checking BGP announcements against IRR data (by resolving the AS-SET object) or RPKI data (ROA objects or a validated cache). It is also common to check the announcements against “bogons” or “martians” (IP prefixes as defined in RFC1918, RFC5735, and RFC6598; ASNs in the AS-PATH as defined by RFC5398, RFC6793, RFC6996, RFC7300, RFC7607).

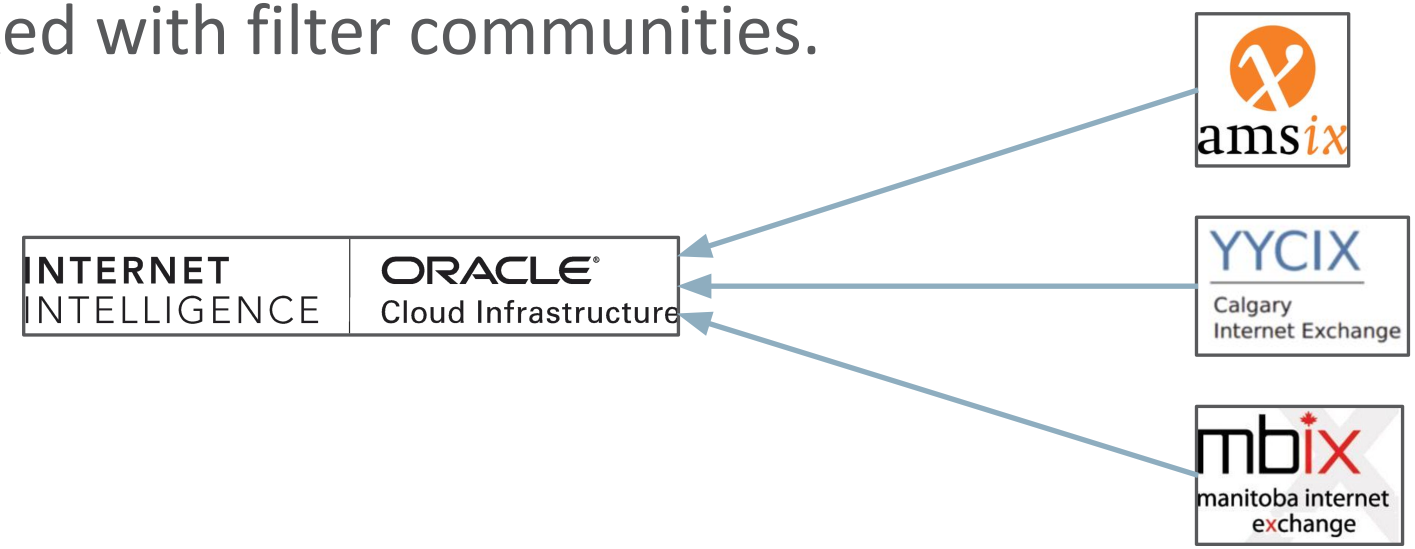
- Measurement could offer:
  - Technical verification of MANRS IXP compliance
  - Feedback for IXPs to ensure proper filtering

- To do it we’ll need route collection from IXP route servers!

ORACLE

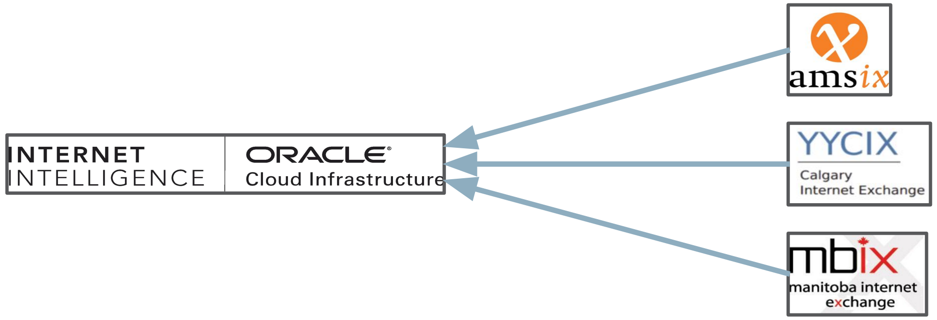
Route collection from IXP route servers.

- Began as an exploratory project by an intern last summer.
- Using direct peering with route servers at three IXPs.
  - *Thank you YYCIX, AMSIX, MBIX*
- Configured as unfiltered sessions - we receive “filtered” routes marked with filter communities.



Route collection from IXP route servers.

- Unfiltered sessions via direct peering allow us to positively confirm invalid announcements are filtered when received.



- However direct peering not easy with most IXPs.
  - Route servers typically not connected to internet.
- And we need **scale** to make an impact.
  - Enter …

PCH
Packet Clearing House

Route collection from IXP route servers.

- PCH publishes MRT files collected at from 180 IXPs around the world.
- These MRT files include PCH’s peering sessions at these IXPs including with the route servers.
- PCH publishes 1 MRT file **every minute** for each IXP (**1440 files/day/IXP**).
- We’re downloading quarter million files per day to learn about routes passed from these IXP route servers.
- The PCH MRT data offers **filtered** sessions with route servers.
  - We sacrifice the unfiltered sessions of direct peering for scale.
  - *Thank you PCH!*

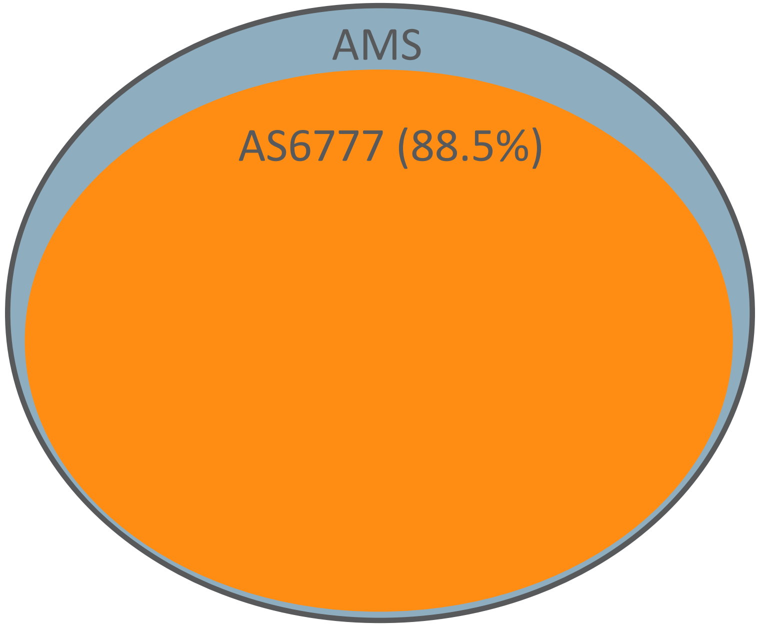
Route collection from IXP route servers.

- What can we do with route collection from 180 IXP route servers?
  - Can’t positively confirm what was filtered (filtered view)
  - But ***can*** flag anything that didn’t get filtered ***but should have***
- Things like
  - RPKI invalids
  - IRR invalids
  - Bogons
- Also things like
  - Spamhaus Droplists *(Not MANRS requirement)*

ORACLE

PCH IXP Data at a glance

- On a recent day 4,597,048 BGP messages collected by PCH at AMSIX.*
  - 4,068,347 (88.5%) were from AS6777 session (AMSIX route server).
  - Next two busiest peering sessions at AMSIX were AS6762 (105,335), AS6461 (79,887).
  - 195 different peering sessions in total.
  - For AS6777, 112,851 unique prefix/origin pairs checked.
    - 2 (0.0018%) were RPKI invalid
  - But AS6777 filters RPKI invalids ...

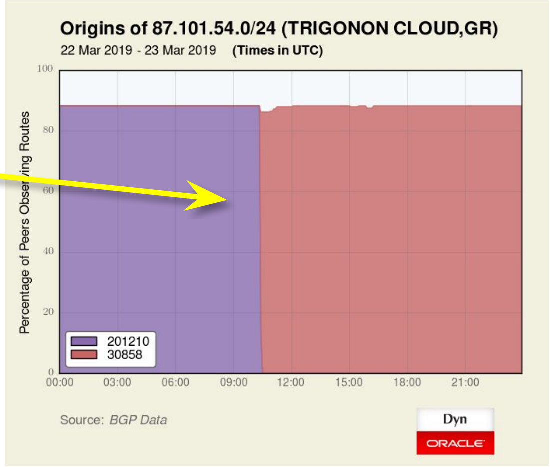


- * 18 March 2019

Why would an RPKI invalid make it through?

```
TIME: 03/22/19 1:27:45 UTC
TYPE: BGP4MP/MESSAGE/Update
FROM: 80.249.209.0 AS6777
ORIGIN: IGP
ASPATH: 57463 8280 201210
NEXT_HOP: 80.249.210.80
COMMUNITY: 0:12989 0:35415 0:62041 12:1089 6777:65011
6777:65012 57463:65000 64700:8280 65400:0 65400:65400
ANNOUNCE
        87.101.54.0/24
```

- BGP origin changed at 10:24:15 UTC on 22 March.
- RIPE NCC RPKI Root (origin also changed)
  - 2019-03-22-ripe-rpki: AS201210 87.101.54.0/24
  - 2019-03-23-ripe-rpki: AS30858 87.101.54.0/24
- Announcement was valid for previous ROA when message sent
- Then invalid for new ROA when message was evaluated.

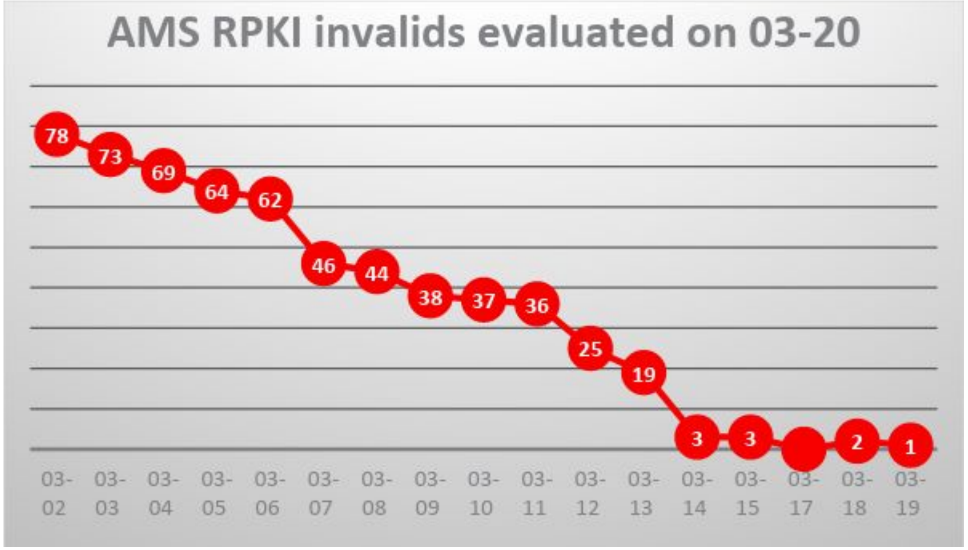


ORACLE

Why would an RPKI invalid make it through?

Message passed by route server $\xrightarrow{t_1}$ PCH collects and archives message $\xrightarrow{t_2}$ We download MRT files from PCH $\xrightarrow{t_3}$ Daily evaluation of messages

- t_1 can be days and t_2 can be up to 24hrs.
- Without time-indexed RPKI validation, some false positives are inevitable.
- Gets worse the farther back you look ------->
- Measuring stick is changing, creating artifacts.
- *When* you validate is as important as *what* you validate!



Where do we see RPKI invalids passed by route servers?

- Top ten route servers based on most unique RPKI invalids on 18 March 2019:

IXP-route server	messages	pfx/origins	RPKI invalids
kbp-59613	512970	83947	571
otp-39107	388267	73482	495
nbo-37704	92985	59379	409
yyz-11670	178696	69056	396
lis-12833	94900	42564	294
fra-6695	470689	35899	161
beg-13004	140631	16268	89
espanix-mad-6895	109016	8681	81
sna-19996	71962	13615	78
pao-24115	140515	13940	77

- 159/161 are /32's with 65535:666
  - (Blackhole as DDoS defense)
- Other two are same measurement artifacts as seem with AMSIX

- Lesson learned: Blackhole announcements override RPKI filters.

ORACLE

Quick Digression:
Where do we see 65535:666 passed by route servers?

- Based on PCH data, top IXPs announcing 65535:666 by BGP message count:

DECIX-Frankfurt	230,595
DECIX-Hamberg	1,229
DECIX-New York	1,141
DECIX-Munich	940
Giganet.ua	654
DECIX-Dusseldof	442
Equinix Singapore	393
Singapore IX	160
SwissIX	157
FranceIX-Paris	121
Equinix Palo Alto	112

- Top IPs with 65535:666 at DECIX-Frankfurt by message count:

201081	109.196.95.199/32	ITCOMP sp. z o.o, PL
7788	202.66.37.41/32	LinkAGE Online Limited, HK
345	46.251.227.2/32	webshield.com, DE
344	46.251.227.4/32	webshield.com, DE
344	46.251.227.35/32	webshield.com, DE

- Top /24's with 65535:666 at DECIX-Frankfurt by message count:

201081	109.196.95.0/24	ITCOMP sp. z o.o, PL
9813	46.251.227.0/24	webshield.com, DE
7788	202.66.37.0/24	LinkAGE Online Limited, HK
188	89.238.5.0/24	Spoldzielnia Mieszkaniowa G
180	109.230.195.0/24	active-servers.com, DE

Based on PCH MRT data collected during 13-19 March

Direct peering w/ route server

just tagging things

LOCATION
Amsterdam

Exchange Services Technical About

Optionally, we can offer the following peering mode in case you really need an unfiltered BGP feed (e.g. for research purposes":

- Peering mode 'Just tagging':
  By selecting this not recommended option, no filtering is applied to announced prefixes. That functionality is helpful for research institutes who want to receive all information or organisations who want to apply their own BGP policies. However, any prefixes that are not filtered will be tagged by using standard BGP communities based on the following criteria (communities are given in the parentheses).
  - Prefix with ROA status: VALID (6777:65012)
  - Prefix with ROA status: INVALID (6777:65022)
  - Prefix with ROA status: UNKNOWN (6777:65023)
  - Prefix present in AS's announced AS/AS-SET (6777:65011)
  - Prefix not present in AS's announced AS/AS-SET (6777:65021)

Direct peering w/ route server

just tagging things

- With *unfiltered* direct peering with route servers we can see what is getting filtered.
  - 'Just tagging' mode, not PCH data
- Some basic statistics:

AMSIX route server (AS6777)

- Total messages
- Prefix with ROA status: VALID (6777:65012)
- Prefix with ROA status: INVALID (6777:65022)
- Prefix with ROA status: UNKNOWN (6777:65023)
- Prefix present in AS's announced AS/AS-SET (6777:65011)
- Prefix not present in AS's announced AS/AS-SET (6777:65021)

ORACLE

Direct peering w/ route server

just tagging things

- With *unfiltered* direct peering with route servers we can see what is getting filtered.
  - 'Just tagging' mode, not PCH data
- Some basic statistics:

AMSIX route server (AS6777)	(13-19)-Mar	
o Total messages	46,851,286	
o Prefix with ROA status: VALID (6777:65012)	5,539,189	11.8%
o Prefix with ROA status: INVALID (6777:65022)	286,860	0.6%
o Prefix with ROA status: UNKNOWN (6777:65023)	37,234,319	79.5%
o Prefix present in AS's announced AS/AS-SET (6777:65011)	21,314,416	45.5%
o Prefix not present in AS's announced AS/AS-SET (6777:65021)	21,751,762	46.4%

ORACLE

Direct peering w/ route server

- Far more prefixes filtered by IRR and AS-Set definitions than RPKI invalids
- Very few instances of bogon prefixes or ASNs, or presence of transit-free ASNs
- Unfiltered view enables greater insight into filtering behavior, but filtered view via PCH helpful to know what *really* gets through.
  - PCH data was essential for identifying bug in YYCIX filtering: RPKI invalid messages were being tagged, but only after passing duplicate message unfiltered.

just tagging things

Direct peering w/ route server

- What are the filtered RPKI invalids at AMSIX? (tagged with 6777:65022)
- 13-19 March 2019: 285157 messages, 1,487 unique prefixes

Count	Prefix	Org
11078	45.115.113.0/24	Drik ICT Ltd, BD
2452	124.81.207.0/24	Indosat Mega Media, ID
2452	124.81.206.0/24	Indosat Mega Media, ID
2451	114.57.35.0/24	Indosat Mega Media, ID
2451	114.57.34.0/24	Indosat Mega Media, ID
2450	114.57.32.0/24	Indosat Mega Media, ID
445	139.5.159.0/24	SKYTEL ISP, LA
402	202.56.4.0/24	Grameenphone Ltd, BD
397	124.106.96.0/20	IPG, PH
396	202.56.5.0/24	Grameenphone Ltd, BD

Count	Prefixes	Peer	Org
53377	320	6939	Hurricane Electric
36792	130	4761	Indosat
24992	79	9498	Bharti Airtel
23940	122	45430	AWN-IIG, TH
22064	112	38158	PT Cyberindo Aditama, ID
18665	110	29049	Delta Telecom, AZ
16071	79	6866	Cyprus Telecom Authority
9446	47	45489	Sri Lanka Telecom
7684	20	14127	Iland Internet Solutions, US
4728	12	37497	Network Platforms LTD, ZA

These 10 peers accounted for 76.4% of messages marked as RPKI invalid by the AMSIX route server.

ORACLE

Direct peering w/ route server

- What are the filtered IRR/AS-Set invalids at AMSIX? (tagged w/ 6777:65021)
- 13-19 March 2019: 21,750,585 messages, 118,863 unique prefixes

Count	Prefix	Org
13273	143.255.0.0/22	Dupont & Schwanke Ltd, BR
11078	*45.115.113.0/24*	*Drik ICT Ltd, BD*
4297	81.85.191.0/24	Globecomm Europe BV, NL
2452	*124.81.207.0/24*	*Indosat Mega Media, ID*
2452	*124.81.206.0/24*	*Indosat Mega Media, ID*
2451	*114.57.35.0/24*	*Indosat Mega Media, ID*
2451	*114.57.34.0/24*	*Indosat Mega Media, ID*
2450	*114.57.32.0/24*	*Indosat Mega Media, ID*
2371	103.73.235.0/24	BPJS Ketenagakerjaan, ID
2313	179.127.207.0/24	Serra Net Telecomunicações, BR

Prefixes in italics also in RPKI invalid top ten.

Count	Peer	Org
3771822	9498	Bharti Airtel
2852499	6939	Hurricane Electric
1780176	9318	SK Telecom
1741590	20485	TransTelekom, RU
856832	7713	Indosat
621924	29049	Delta Telecom, AZ
511031	52320	GlobeNet
497741	37468	Angola Cables
472578	6663	Turk Telekom
389597	37100	Seacom

These 10 peers accounted for 62% of messages marked as IRR invalid by AMSIX route server.

ORACLE

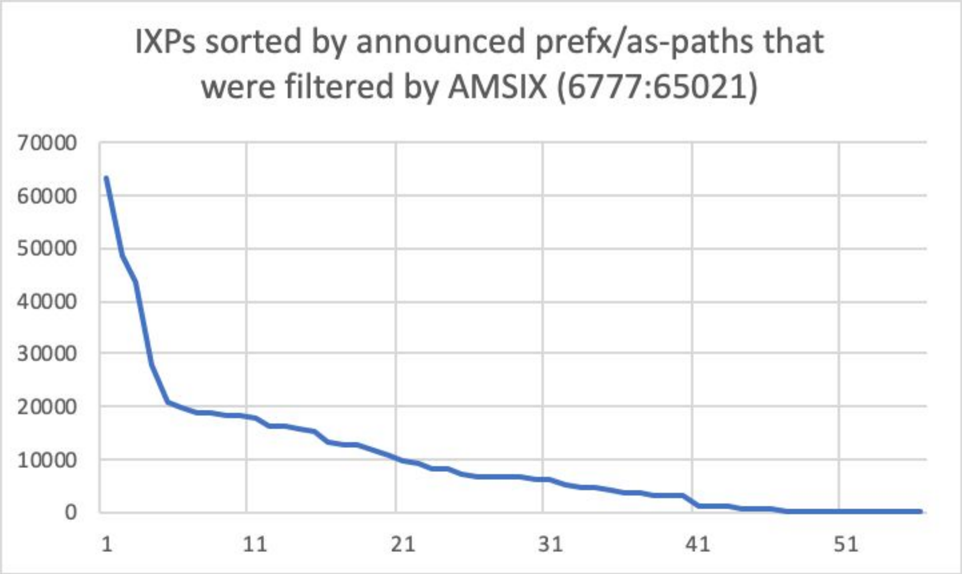
Challenge: Differences in IRR/AS-set filtering at IXPs

- AMSIX filtered messages (marked with 6777:65021) seen at other IXPs:
  - BGP4MP|1552435659|A|80.249.208.255|6777|198.105.254.0/24|6939 36029|IGP|80.249.209.150|0|1428|0:12876 0:12989 0:13335 0:15169 0:16265 0:16276 0:20940 0:22822 0:29073 0:32768 0:32934 0:43350 0:49981 6777:65021 6777:65023|
- Announcement not seen for AMSIX in PCH (filtered), but seen in PCH data for other IXPs:
  - FranceIX-Marseilles: BGP4MP|1552497208|A|37.49.236.251|51706|198.105.254.0/24|6939 36029|IGP|37.49.236.10|0|1335|51706:64602 51706:64650 51706:65011 51706:65023|
  - DECIX-Hamburg: BGP4MP|1552453965|A|80.81.203.252|43252|198.105.254.0/24|6939 36029|IGP|80.81.203.55|0|1508|65101:4081 65102:4000 65103:276 65104:150|
  - ECIX-Munich: BGP4MP|1552435622|A|194.59.190.254|9033|198.105.254.0/24|6939 36029|IGP|194.59.190.69|0|1495|9033:3302 65000:12989 65000:13335 65000:15169 65000:16265 65000:20940 65000:22822 65010:1 65011:1|
  - Japan IX: BGP4MP|1552523969|A|210.171.224.11|23815|198.105.254.0/24|6939 36029|IGP|210.171.224.40|0|1080|23815:0|
  - (and 30+ others)

ORACLE

Challenge: Differences in IRR/AS-set filtering at IXPs

- If we take identical prefix/AS-path pairs filtered at AMSIX and look for them at other IXPs in PCH data, we find:



Top ten IXPs passing routes filtered at AMSIX:

Count	RS ASN	IXP
63240	57463	NetIX
48452	6695	DECIX
43434	24115	Equinix Singapore
27703	51706	France IX
21079	39107	Interlan IX
19954	37195	NAPAfrica
18903	12833	GigaPix
18680	34307	NL-ix
18235	8714	LINX
18164	20717	DECIX

- *Not concluding that one IXP is correct and another is wrong, just highlighting variance among IRR/AS-set filter implementations.*

ORACLE

How to objectively measure MANRS compliance for filtering?

> Validation is usually done by checking BGP announcements against IRR data (by resolving the AS-SET object) or RPKI data (ROA objects or a validated cache). It is also common to check the announcements against "bogons" or "martians" (IP prefixes as defined in RFC1918, RFC5735, and RFC6598; ASNs in the AS-PATH as defined by RFC5398, RFC6793, RFC6996, RFC7300, RFC7607).

- IRR filtering can vary by implementation and data employed.
- RPKI filtering more straightforward, but is less common presently.
- IRR vs RPKI: If either is allowed, then can an IXP be compliant if it scores well for one and not the other? Based on our preliminary analysis, no IXP scores perfect in both.
  - ECIX locations have 0 IRR invalids in PCH data, but have some RPKI invalids.

	Messages	Pfx/Origins	RPKI inval	IRR inval
ECIX-Munich	142474	43317	157	0
ECIX-Hamberg	35133	4620	10	0
ECIX-Frankfurt	38285	5684	10	0
ECIX-Dusseldorf	36284	4706	10	0
ECIX-Berlin	62771	11328	63	0

ORACLE

How to objectively measure MANRS compliance for filtering?

- So what next? Too early to define criteria for compliance yet.
- Next step: publish filter statistics based on PCH data.
  - Updated periodically (weekly?)
  - Include invalid messages observed to enable scrutiny
- Collect feedback from the IXP community!

MANRS

ORACLE

Big thanks to…

- Job Snijders, NTT Communications
- PCH (Gaël Hernández & Ashley Jones)
- Theo deRaadt, Calgary IX
- Stefan Plug, ECIX
- Theo Baschak, MBIX
- Andrei Robachevsky, ISOC

douglas.madory@oracle.com
@dougmadory

