The Internet Outage on Aug. 25
from the point of view of IX

報告人 魯堯

Original Version From:

Nasato Goto

APIX#16

September 11th 2017

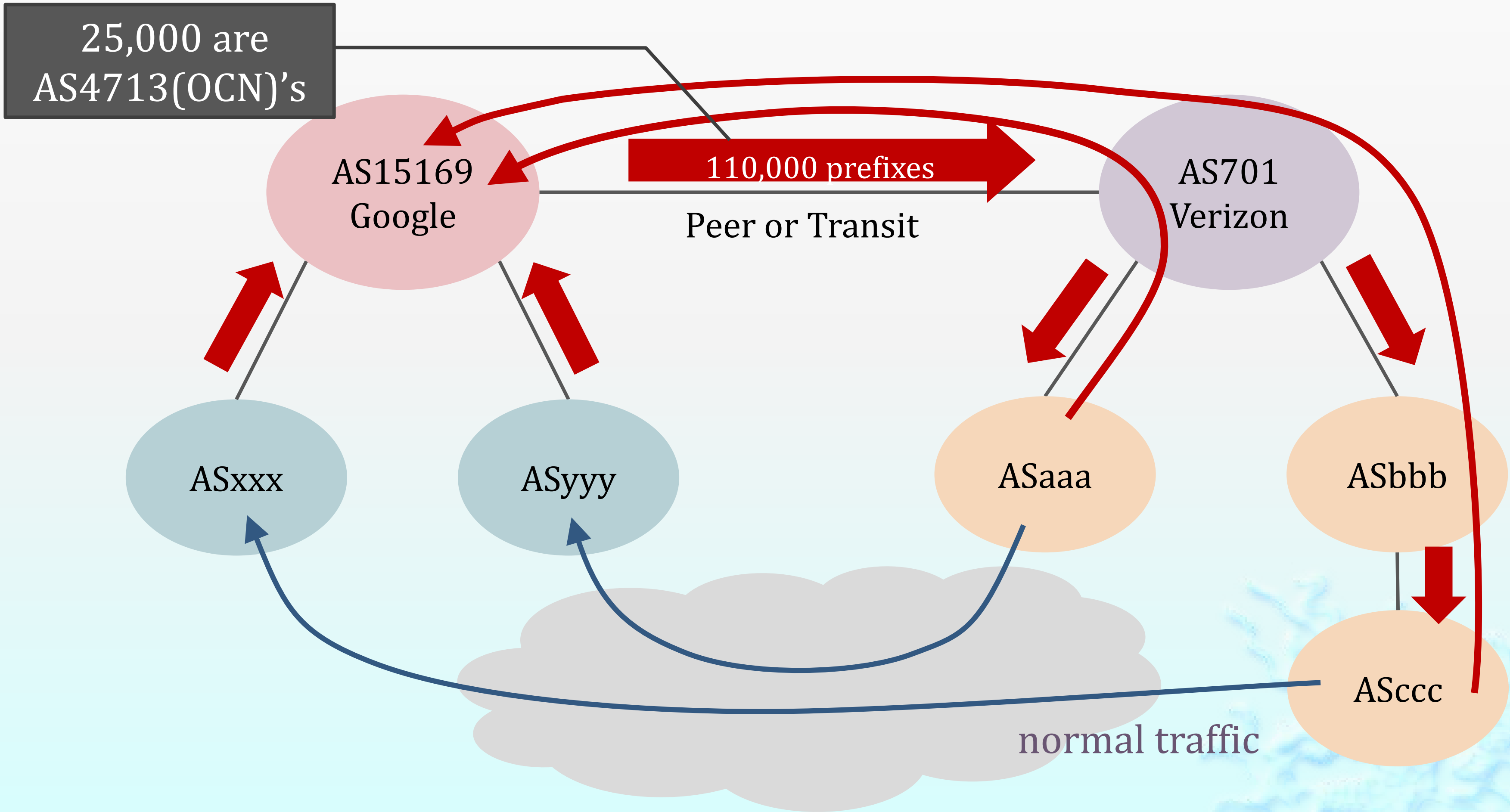
What happened on 08/25?

- Timeline
  - 2017/08/25 12:22 (JST)
    - AS15169 started to announce many IPv4 prefixes, totally 110,000.
      - More specific prefixes were detected at that time.
    - The network failures were detected in Japan.
  - 2017/08/25 12:30 (JST)
    - (AS15169 says) they withdrew the prefixes.
- Main impact of this route leak
  - (1) Unusual traffic forwarding toward AS15169
  - (2) Router performance decrement
- Other influence
  - IX segment hijacking

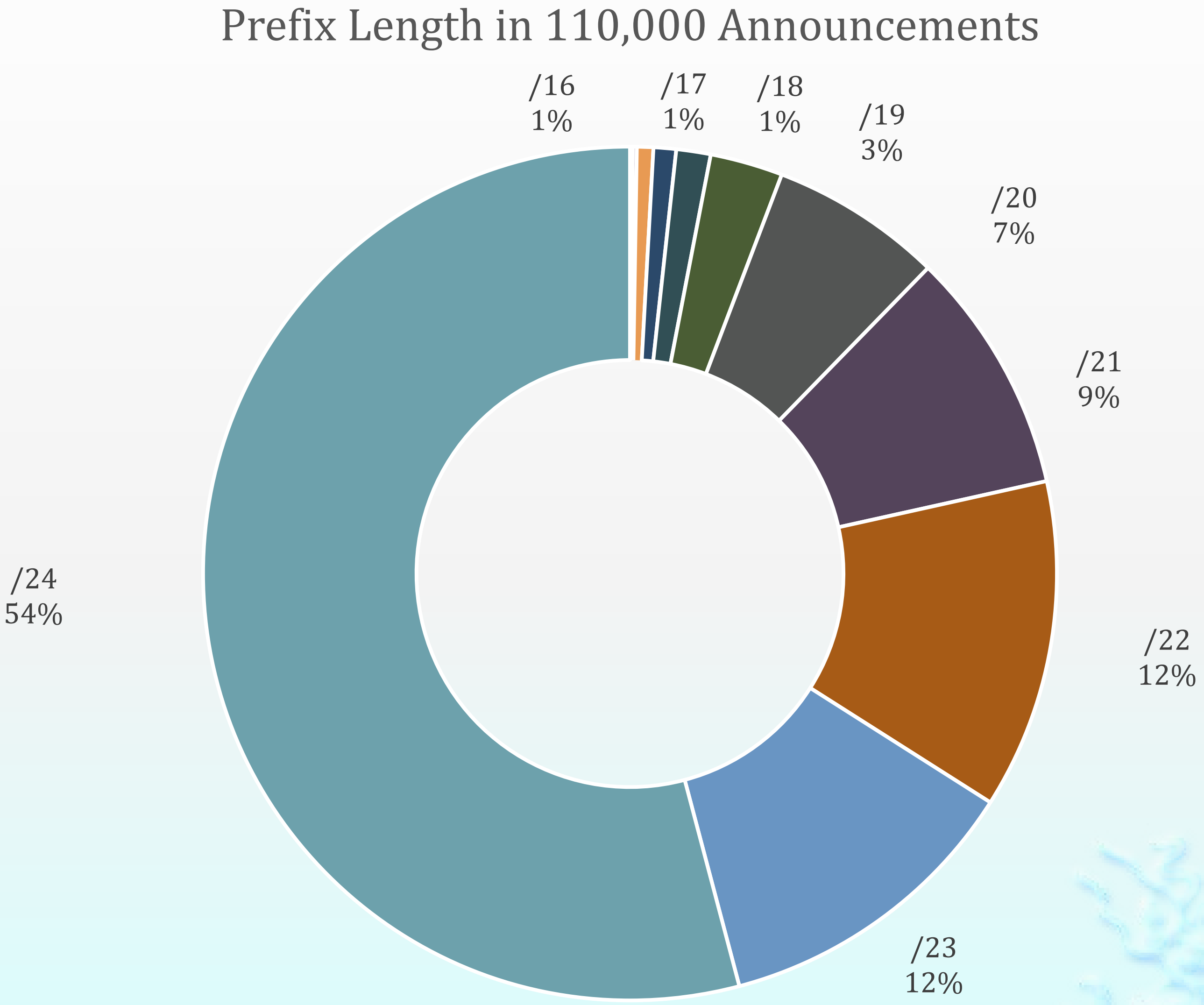
cite from)
http://www.asahi.com/ajw/articles/AJ201708270030.html
https://www.attn.jp/maz/p/t/pdf/20170825-routeleakage.pdf (Japanese)

Influence of route leak (1/2)

- (1) Unusual traffic forwarding
  - According to the more specific prefixes announced by AS15169, traffic flew into AS15169, via AS701.



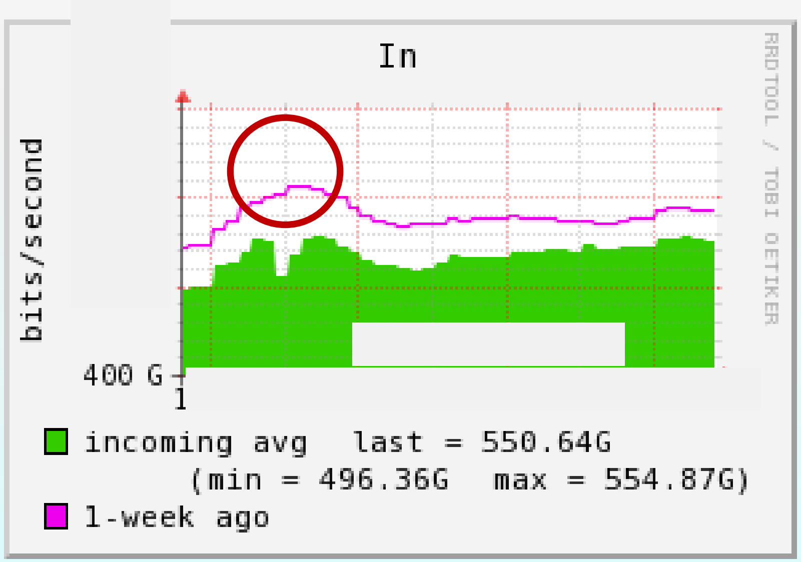
Prefix Length of Leaked Routes

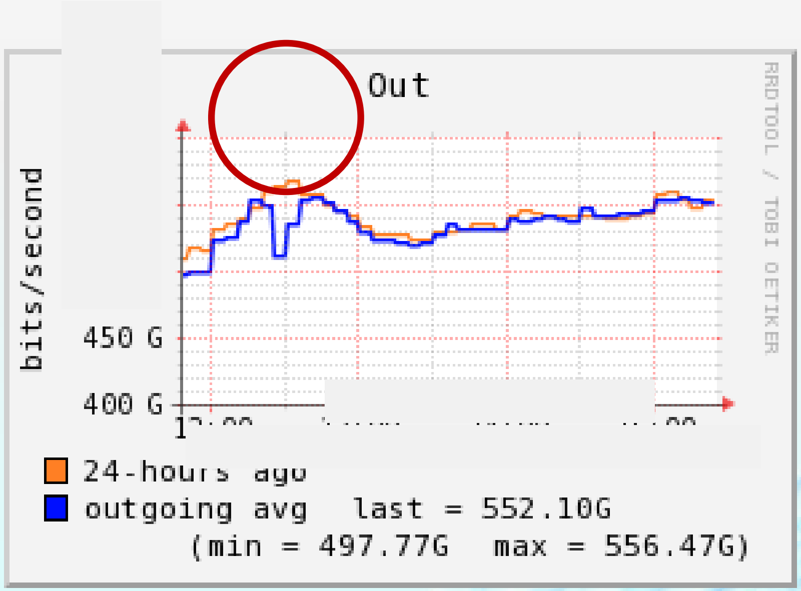


datasource) http://archive.routeviews.org

Influence of route leak (1/2)

- (1) Unusual traffic forwarding
  - Traffic influence was observed in JPNAP.
  - Both of In/Out traffic decrement (encircled red) were seemed to be moved from JPNAP to others, or blackholed.

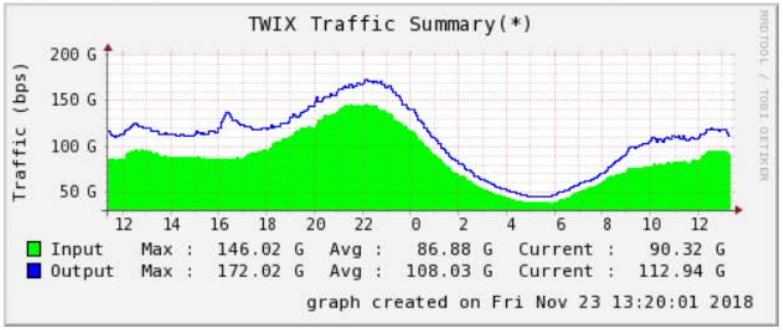




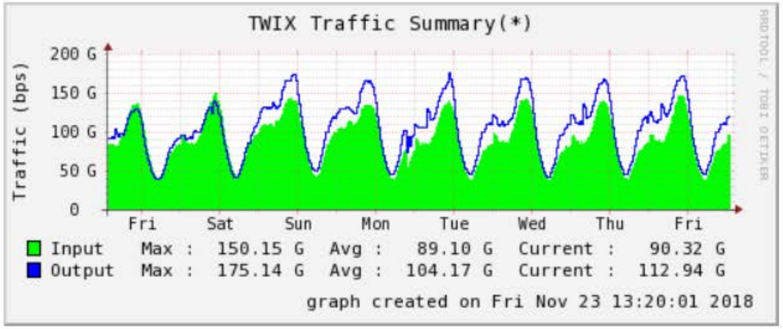
TWIX流量

Total Traffic For TWIX

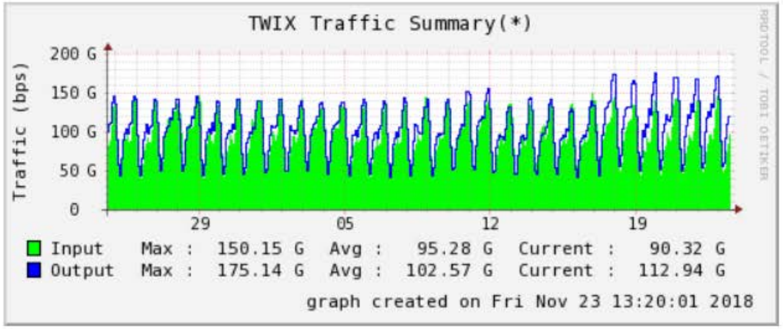
(*)106年4月12日起，TWIX將Public Peering與Private Peering訊務併計為IX訊務總量



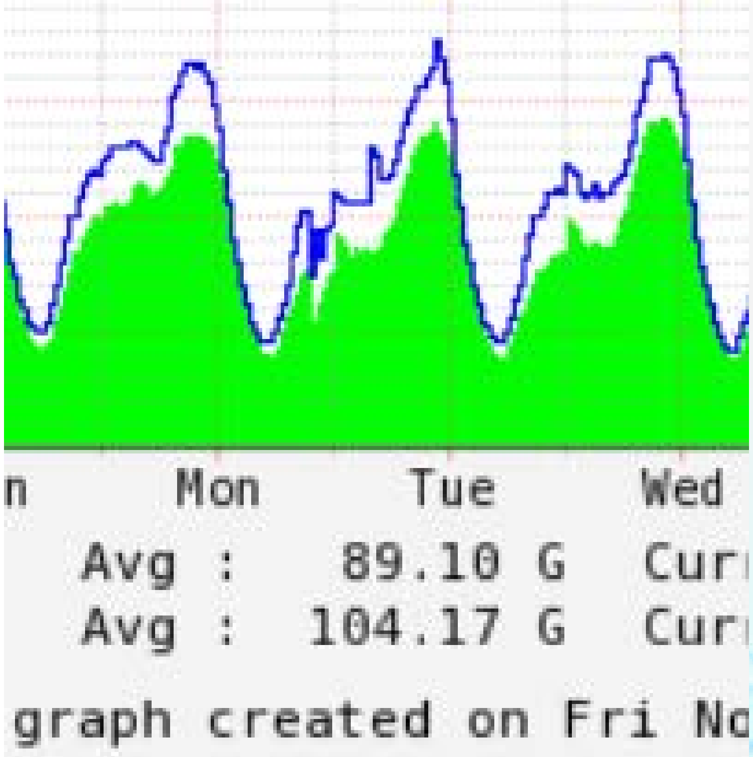
```
🟩 Input    Max :  146.02 G   Avg :   86.88 G   Current :   90.32 G
🟦 Output   Max :  172.02 G   Avg :  108.03 G   Current :  112.94 G
                     graph created on Fri Nov 23 13:20:01 2018
```



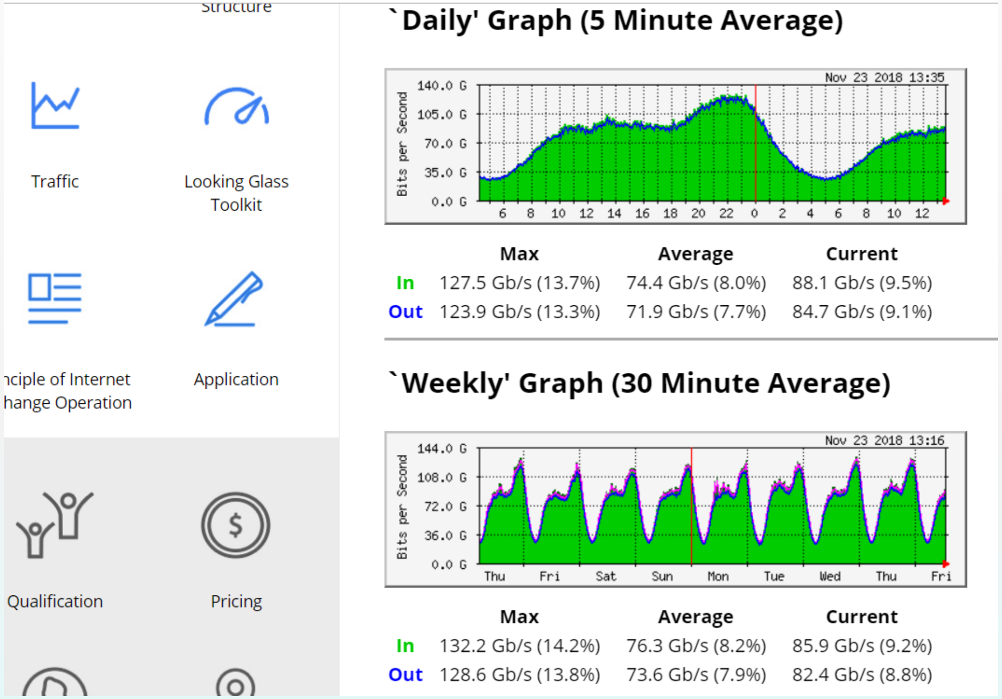
```
🟩 Input    Max :  150.15 G   Avg :   89.10 G   Current :   90.32 G
🟦 Output   Max :  175.14 G   Avg :  104.17 G   Current :  112.94 G
                     graph created on Fri Nov 23 13:20:01 2018
```



```
🟩 Input    Max :  150.15 G   Avg :   95.28 G   Current :   90.32 G
🟦 Output   Max :  175.14 G   Avg :  102.57 G   Current :  112.94 G
                     graph created on Fri Nov 23 13:20:01 2018
```



TPIX

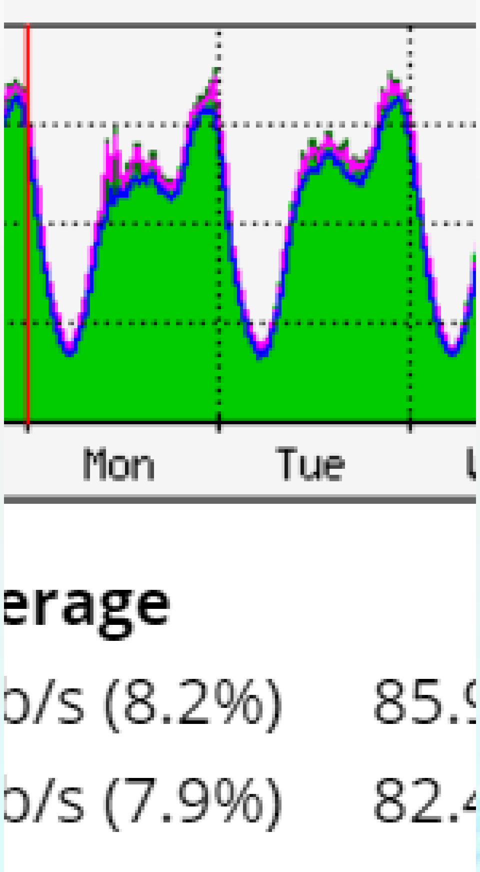


`Daily' Graph (5 Minute Average)

	Max	Average	Current
In	127.5 Gb/s (13.7%)	74.4 Gb/s (8.0%)	88.1 Gb/s (9.5%)
Out	123.9 Gb/s (13.3%)	71.9 Gb/s (7.7%)	84.7 Gb/s (9.1%)

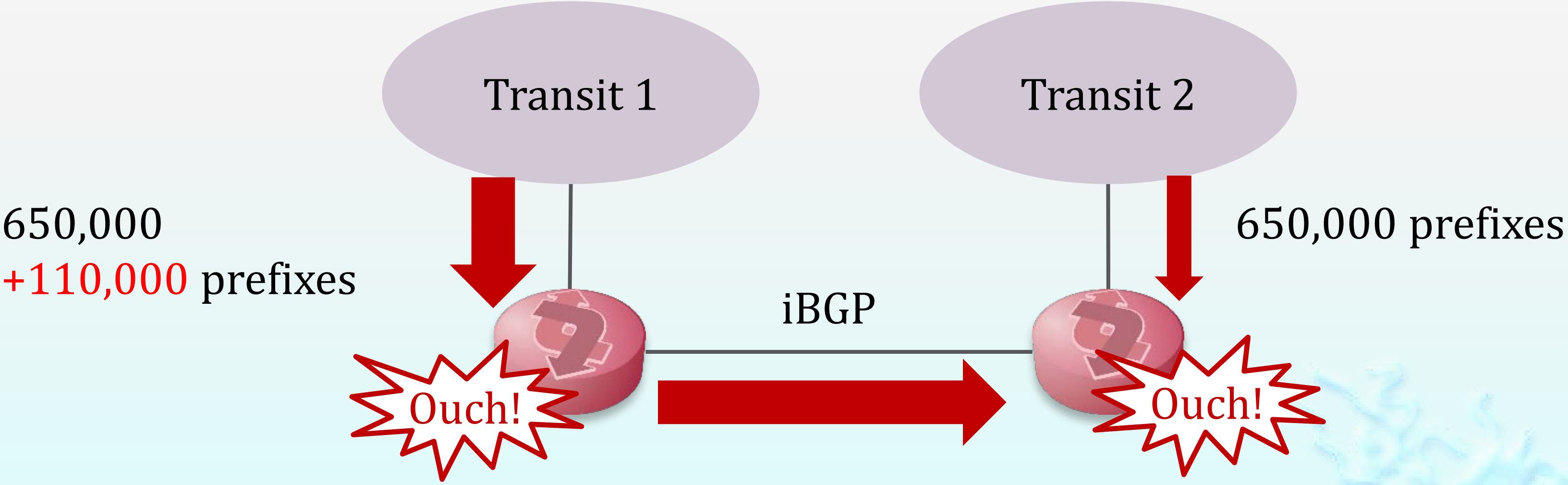
`Weekly' Graph (30 Minute Average)

	Max	Average	Current
In	132.2 Gb/s (14.2%)	76.3 Gb/s (8.2%)	85.9 Gb/s (9.2%)
Out	128.6 Gb/s (13.8%)	73.6 Gb/s (7.9%)	82.4 Gb/s (8.8%)



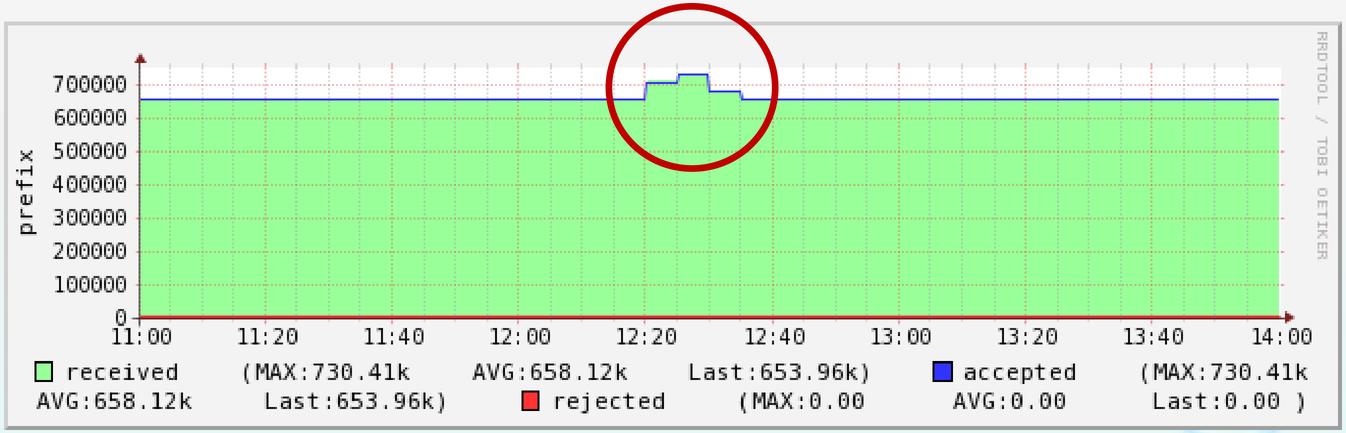
Influence of route leak (2/2)

- (2) Router performance decrement
  - Some Japanese ISPs still use router whose TCAM size is not so big.
  - Because of explosive increase of full route, the TCAM overflowed. This caused performance decrement to the routers.



Example: Influence on Japanese ISP

- transix (AS55391/55392)
  - provides IPv6 Internet service.
  - and also provides IPv4 connectivity over IPv6 as a option service.
- The Backbone router received more than 700,000 prefixes from its transit IIJ (AS2497) at that time.



- But IPv4 traffic was seemed not to be affected.
  - (Guess) This is because the leaked prefixes didn’t include target IP addresses of transix IPv4 traffic.11

IX Segment Hijacking

IXP	Google?	Hijacked?
AMS-IX Hong Kong	No	No
BBIX	Yes	Yes
BDIX	No	No
BKNIX	No	No
CHN-IX	No	No
CNX	No	No
DIX-IE	Yes	Yes
Equinix	Yes	Yes
HKIX	Yes	No
IIX	No	No
IX-Australia	Yes	Yes
JPIX	Yes	Yes

IXP	Google?	Hijacked?
JPNAP	Yes	Yes
KINX	No	No
Megaport	Yes	Yes
MumbaiIX	No	No
MyIX	Yes	Yes
NIXI	No	No
NPIX	No	No
NZIX	?	No
PHOpenIX	?	?
SGIX	Yes	Yes
SOX	Yes	Yes
TPIX	Yes	Yes
VNIX	?	No

- 11 out of 25 APIX member IXs suffered hijack of their IX segment.
- This event might have affected to traffic in IX.

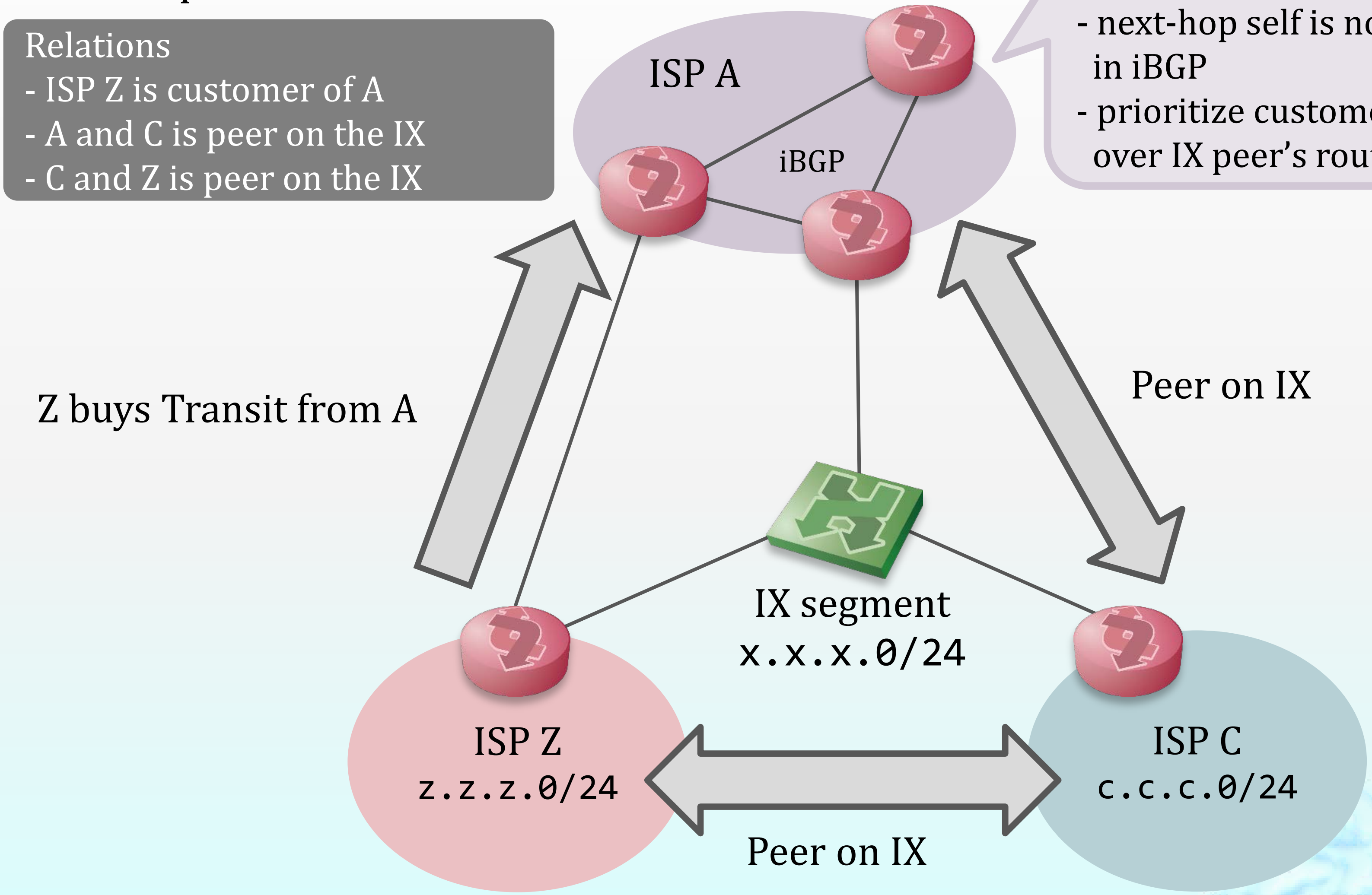
data source) Peering DB, Route Views Archive

Influence of IX Segment Hijacking (0/4)

- Example Conditions

Relations
- ISP Z is customer of A
- A and C is peer on the IX
- C and Z is peer on the IX

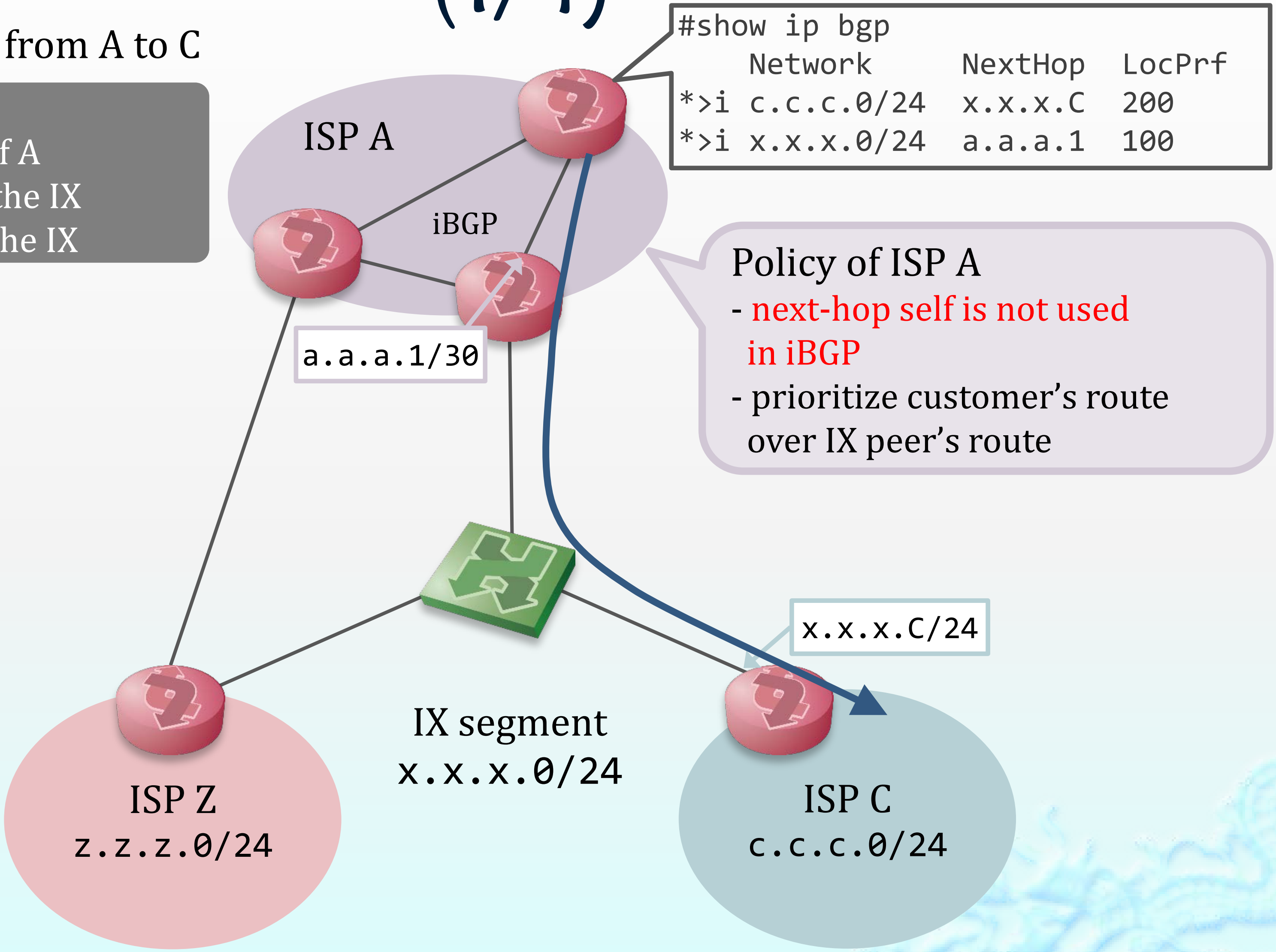
Policy of ISP A
- next-hop self is not used in iBGP
- prioritize customer's route over IX peer's route



Influence of IX Segment Hijacking (1/4)

- Normal Traffic from A to C

Relations
- ISP Z is customer of A
- A and C is peer on the IX
- C and Z is peer on the IX



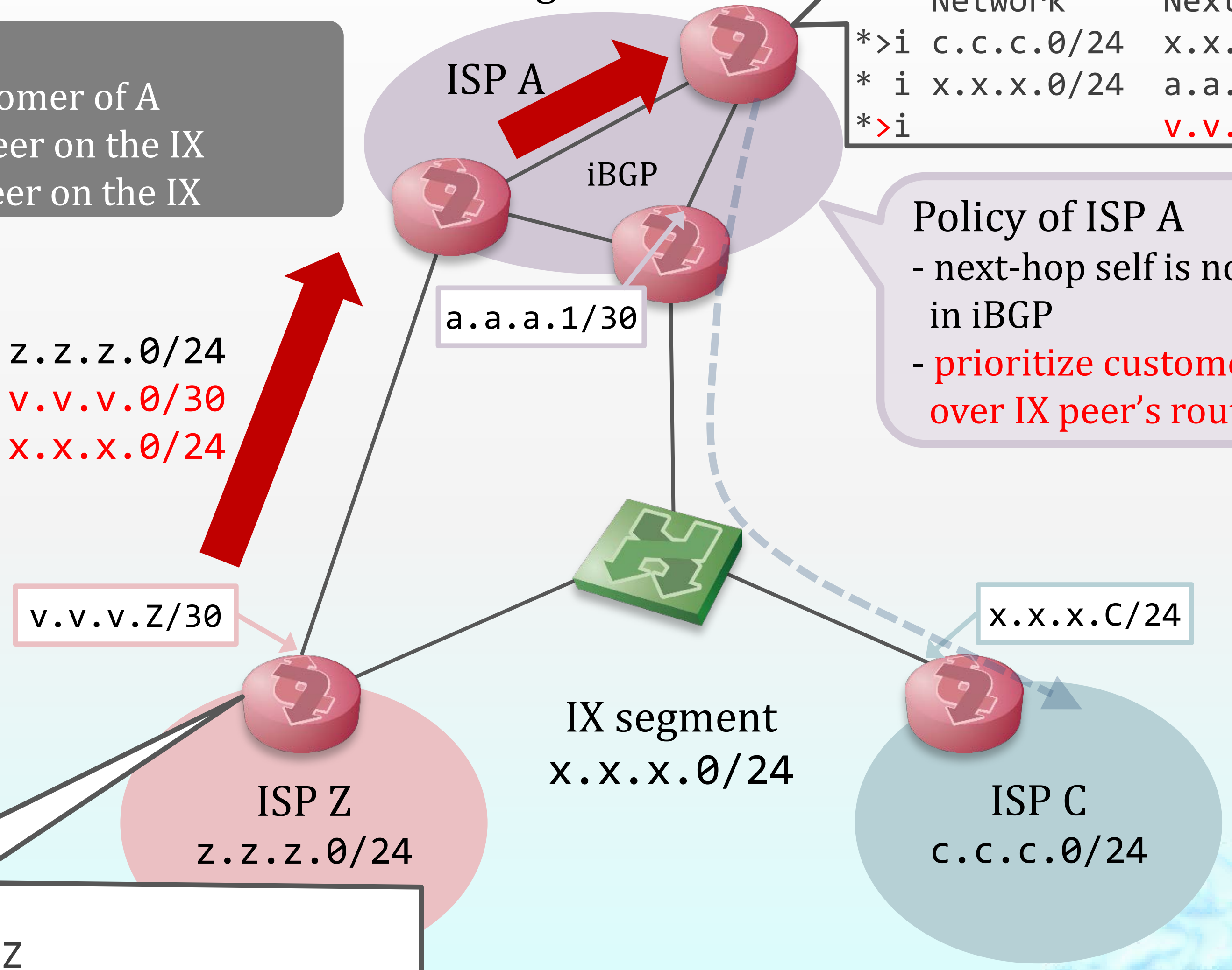
```
#show ip bgp
     Network        NextHop   LocPrf
*>i c.c.c.0/24   x.x.x.C   200
*>i x.x.x.0/24   a.a.a.1   100
```

Policy of ISP A
- next-hop self is not used in iBGP
- prioritize customer's route over IX peer's route

Influence of IX Segment Hijacking (2/4)

- Z starts to announce its connected segment

Relations
- ISP Z is customer of A
- A and C is peer on the IX
- C and Z is peer on the IX



```
#show ip bgp
     Network        NextHop   LocPrf
*>i c.c.c.0/24     x.x.x.C   200
*  i x.x.x.0/24     a.a.a.1   100
*>i                v.v.v.Z   300
```

Policy of ISP A
- next-hop self is not used in iBGP
- prioritize customer's route over IX peer's route

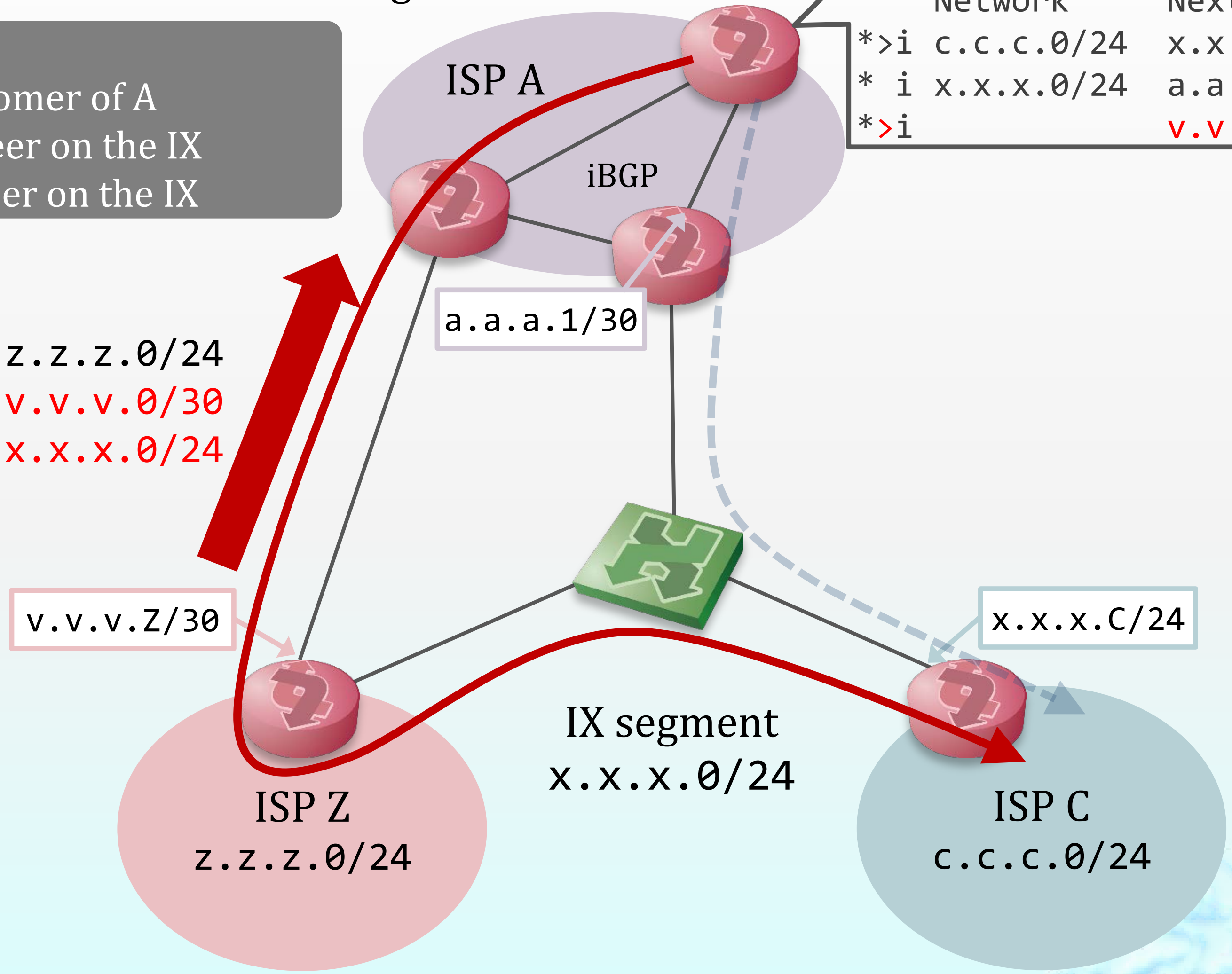
```
# conf t
router bgp Z
 redistribute connected
 neighbor v.v.v.A remote-as A
 neighbor x.x.x.C remote-as C
```

Influence of IX Segment Hijacking (3/4)

- Traffic from A to C flow through Z

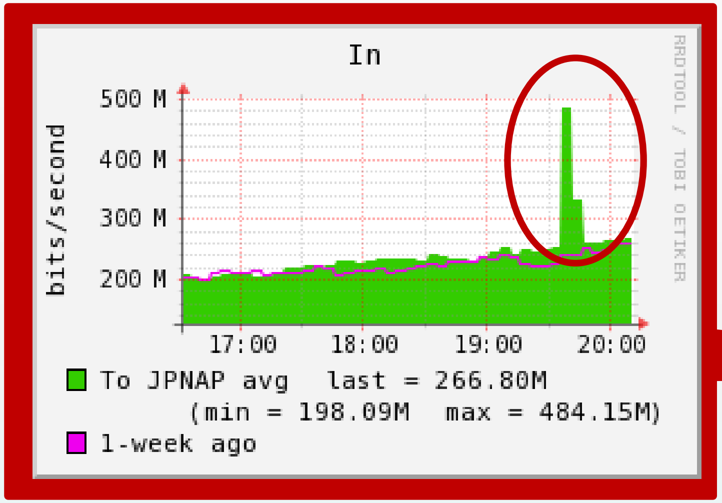
Relations
- ISP Z is customer of A
- A and C is peer on the IX
- C and Z is peer on the IX

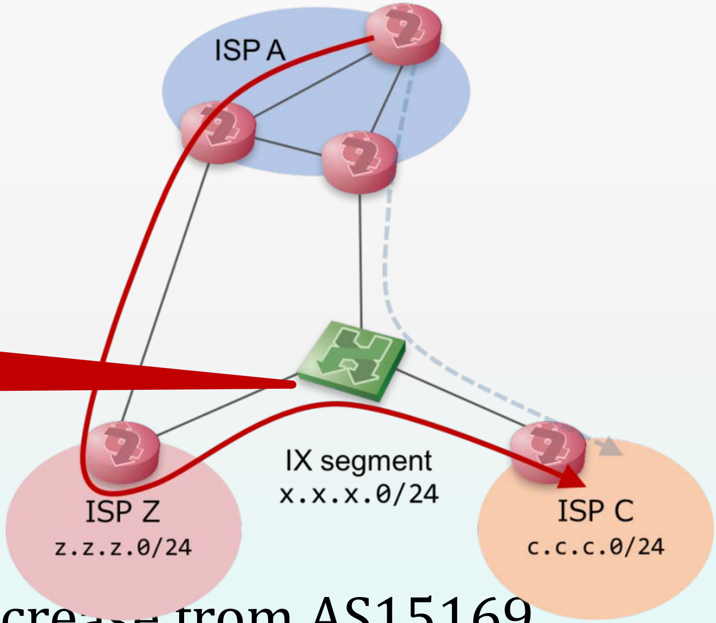
```
#show ip bgp
    Network        NextHop    LocPrf
*>i c.c.c.0/24     x.x.x.C    200
* i x.x.x.0/24     a.a.a.1    100
*>i                v.v.v.Z    300
```



Influence of IX Segment Hijacking (4/4)

- Past example at JPNAP
  - When a customer leaked our IX segment, the traffic graph of the customer showed spike due to influence of the hijack.





- This time
  - We didn't observe the traffic increase from AS15169.
- Therefore, in JPNAP, we had no hijacking influence on our traffic.

Q & A

- Max-prefix-limit configuration on eBGP routers to ISPs
- Better Router
- Think 3 times before you move
- And???