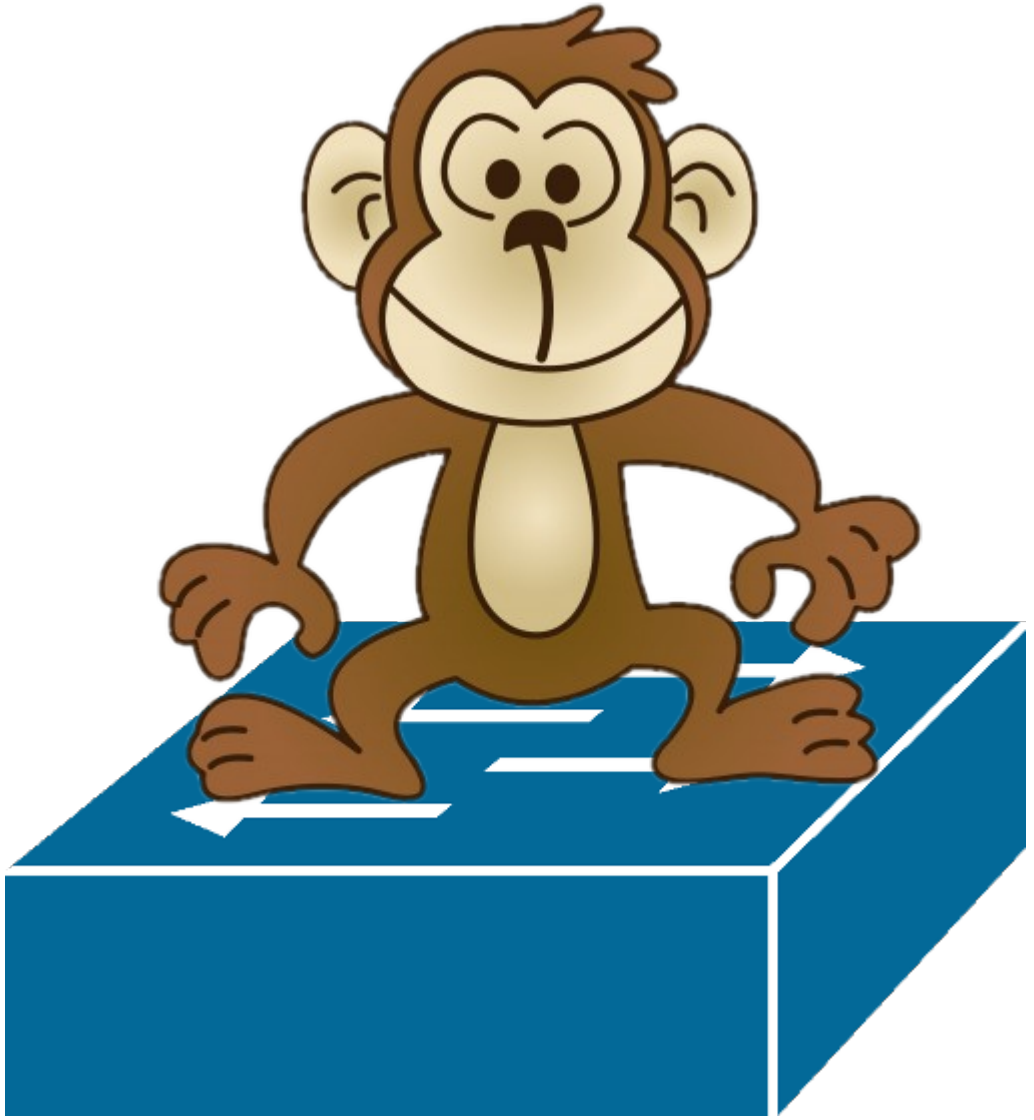
Monkeying Around on the APE



Mike Jager

APNIC 30
Gold Coast, Australia
August 2010

- Auckland Peering Exchange (APE), New Zealand
- Attached laptop to new port

```
default via fe80::20c:cfff:febd:d7ca dev eth0
default via fe80::21e:beff:fec0:c21a dev eth0
default via fe80::20b:5fff:fefd:cbb1 dev eth0
default via fe80::207:b3ff:fe5e:5221 dev eth0
```

- Hey look, internet!

Me

- Senior Network Engineer
  - Web Drive, Auckland, New Zealand
- We do the content bit
- Web hosting, domain names, servers

When connecting your network to a shared layer 2 network such as an Internet Exchange Point, make sure you take appropriate steps to protect your network against misuse

IXP fundamentals

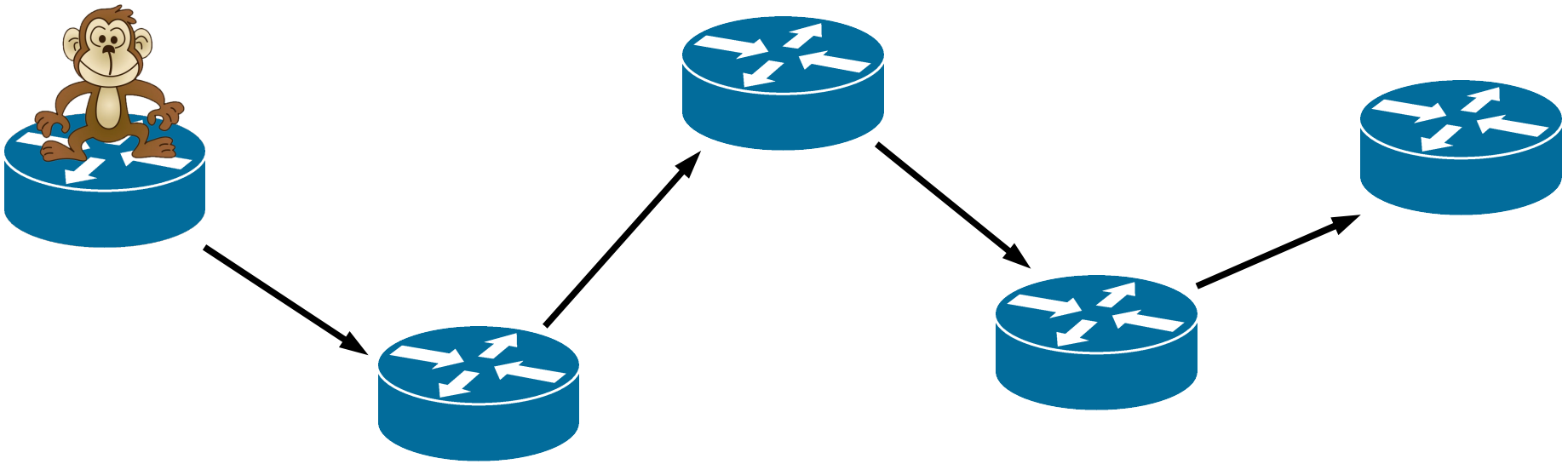
- Shared Layer 2 network
- IXP operator assigns IP addresses to members
- Members stand up BGP sessions
  - between each other
  - to Route Servers (run by IXP operator)
- Routes exchanged, packets flow

IXP fundamentals

- Even bilateral peering across IXP fabric is much simpler than standing up one circuit per peer
- Helps reduce requirement for transit (save $$)
- Keeps local traffic local
  - reduced latency, jitter
  - increased bandwidth between peers?
- Easy access by you to other IXP members
- Easy access by other IXP members to you

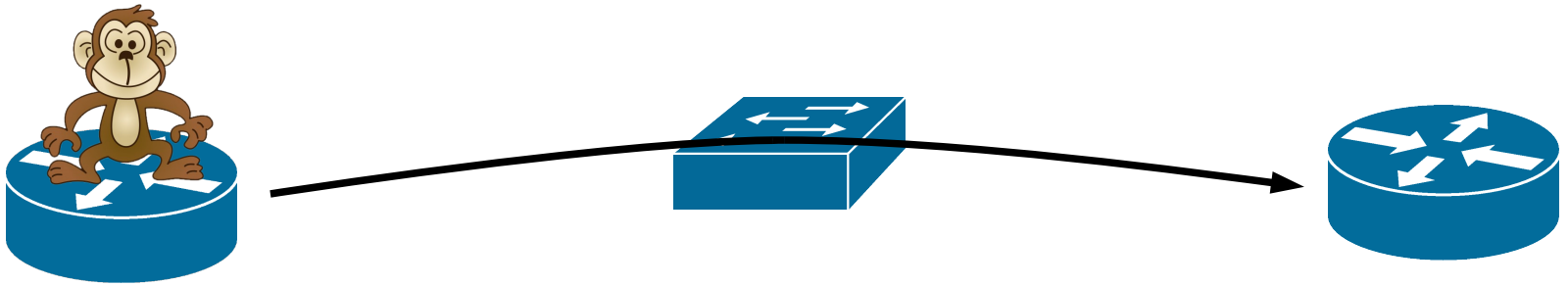
IXP = cheap access to your routers

- Transit: dedicated, private, point-to-point circuit
- To send arbitrary packets to your network via transit, intermediary routers must forward IP packets across the internet to you



IXP = cheap access to your routers

- IXPs enable someone to throw any packet directly at the border of your network



- Any ethernet frame with a destination MAC of your router will end up at its IXP-facing interface
- Routers attached to shared layer 2 networks are more vulnerable to receiving malicious packets

Packet dumping an IXP

- Additionally, any broadcast/multicast packets your router is sending can be detected
- IXP is a switched ethernet
- Assuming exchange traffic only:
  - unicast traffic destined for packet dumping machine
  - non-unicast traffic required for IXP operation
    - broadcast ARP for IXP IPv4 addresses
    - multicast IPv6 NS for IXP IPv6 addresses
  - multicast?

Packet dumping the APE

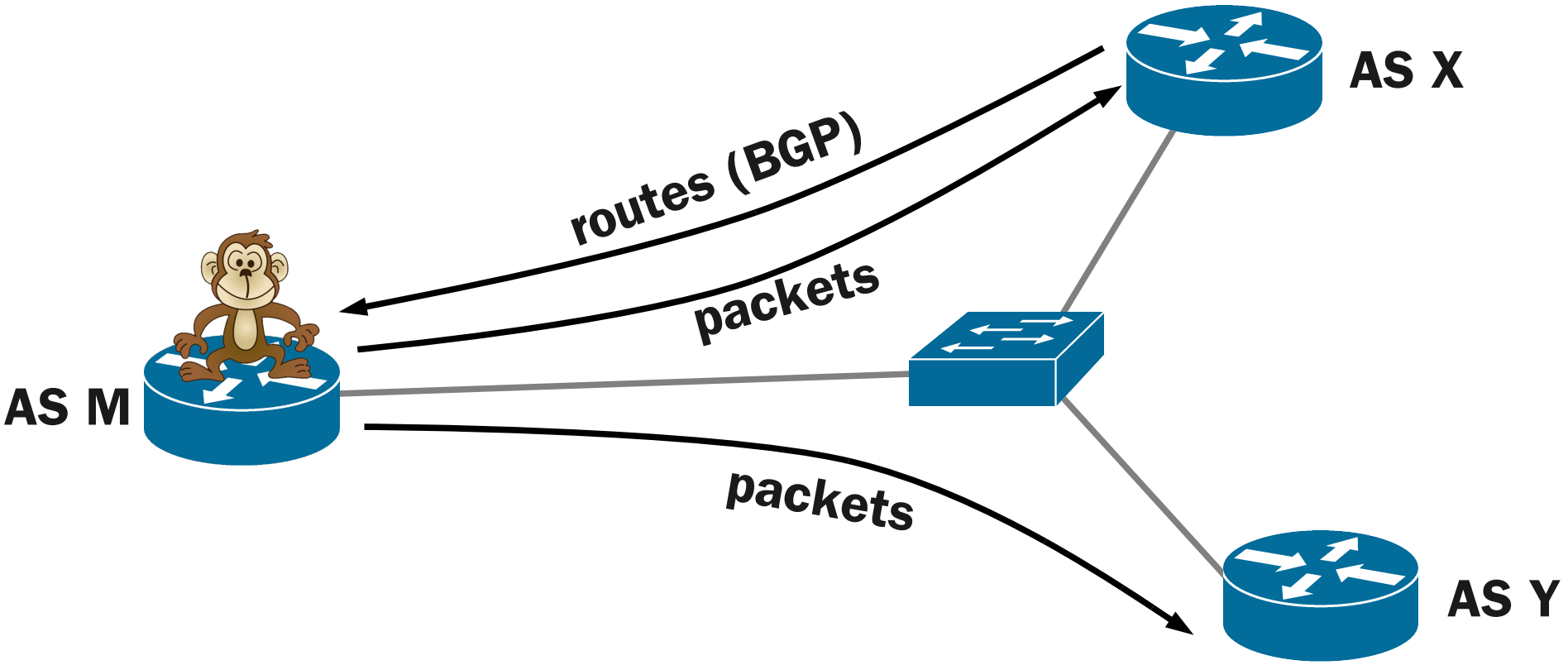
- ARP, and lots of it
  - lots to non-APE address space
    - bilateral peering sessions not using APE /24?
    - selling transit across APE?
- DHCP
- IPv6 RAs
- PIM/IGMP/MLD
- DECNET MOP
- ~~OSPF (?!)~~

Borrowing other networks

- Packets forwarded according to best path from routing table
- Next-hop address found, MAC discovered, frame put on wire
- What if someone ignores the routing table, and chooses their own destination MAC?
  - (or gets creative setting the next-hop on prefixes learned from elsewhere)

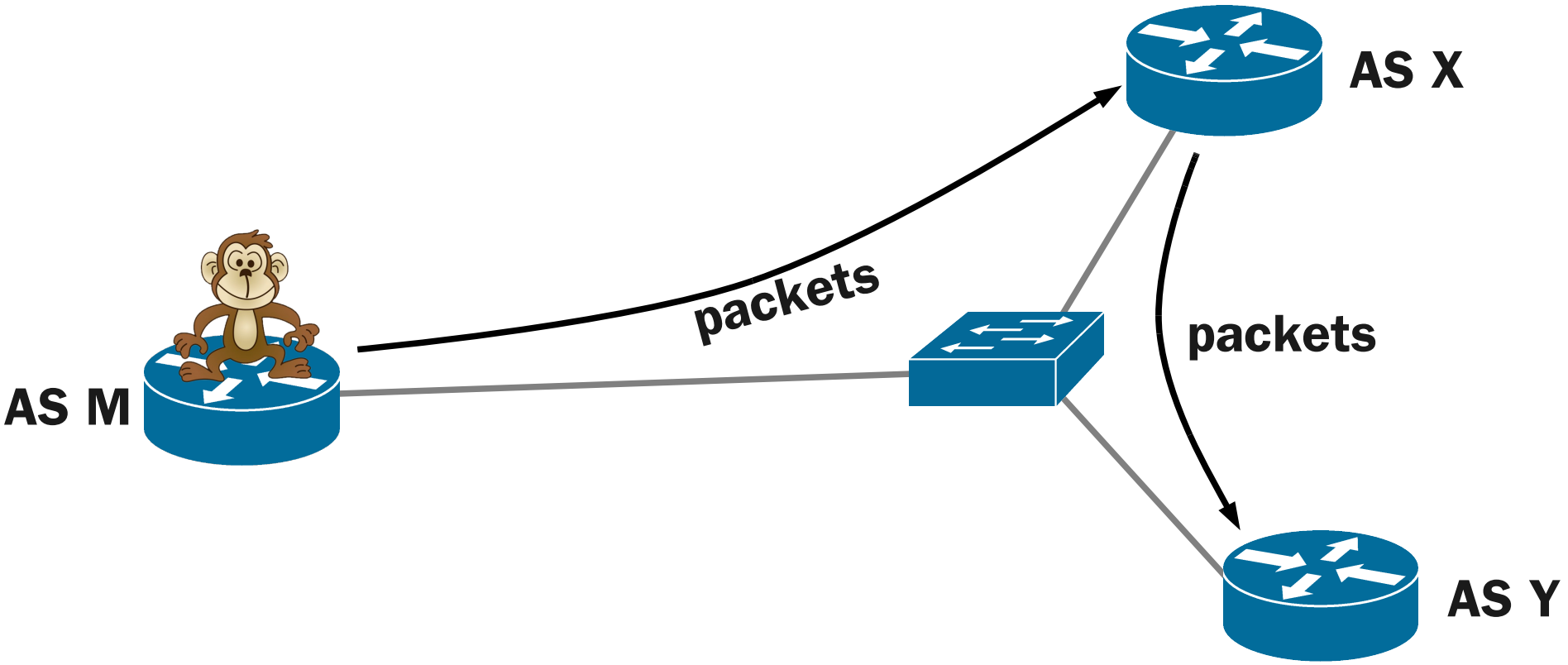
Borrowing other networks

- Directly into a AS Y that wont peer with AS M



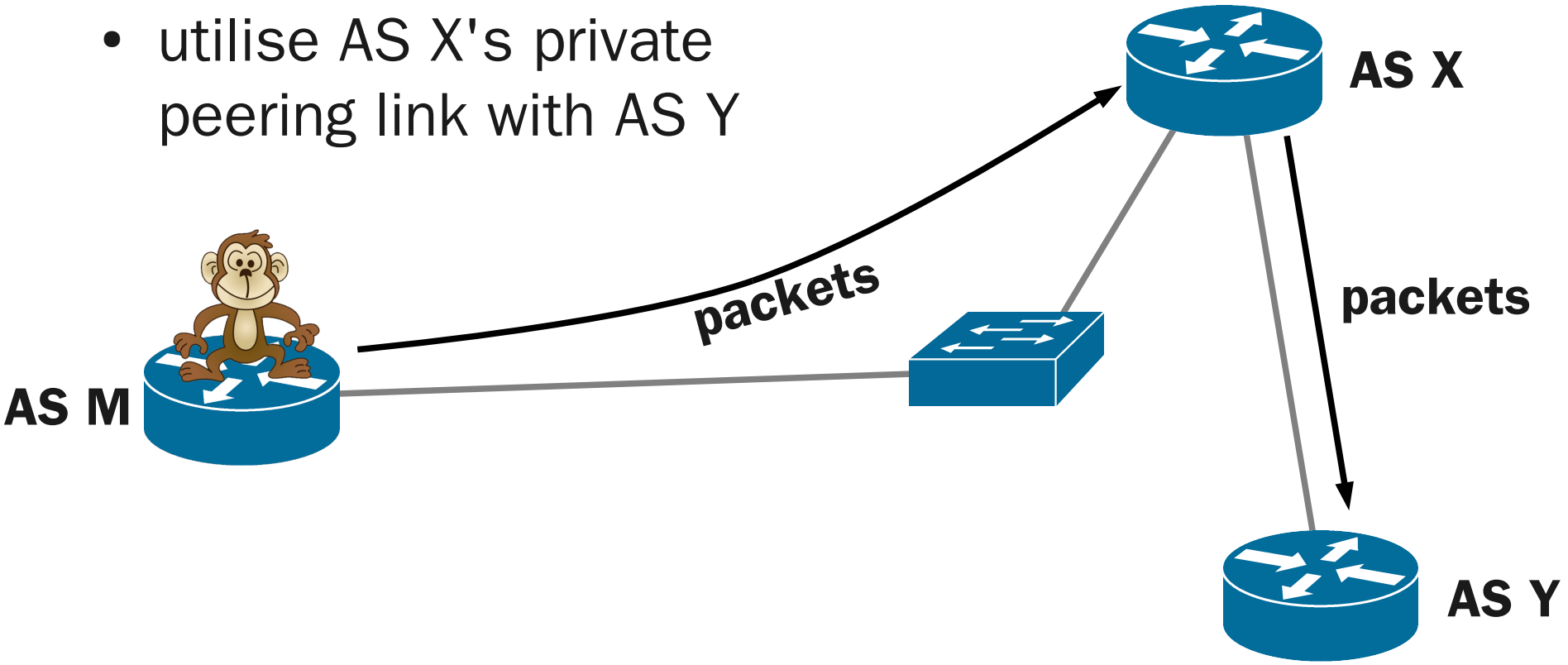
Borrowing other networks

- Into AS X to reach AS Y that wont peer with AS M



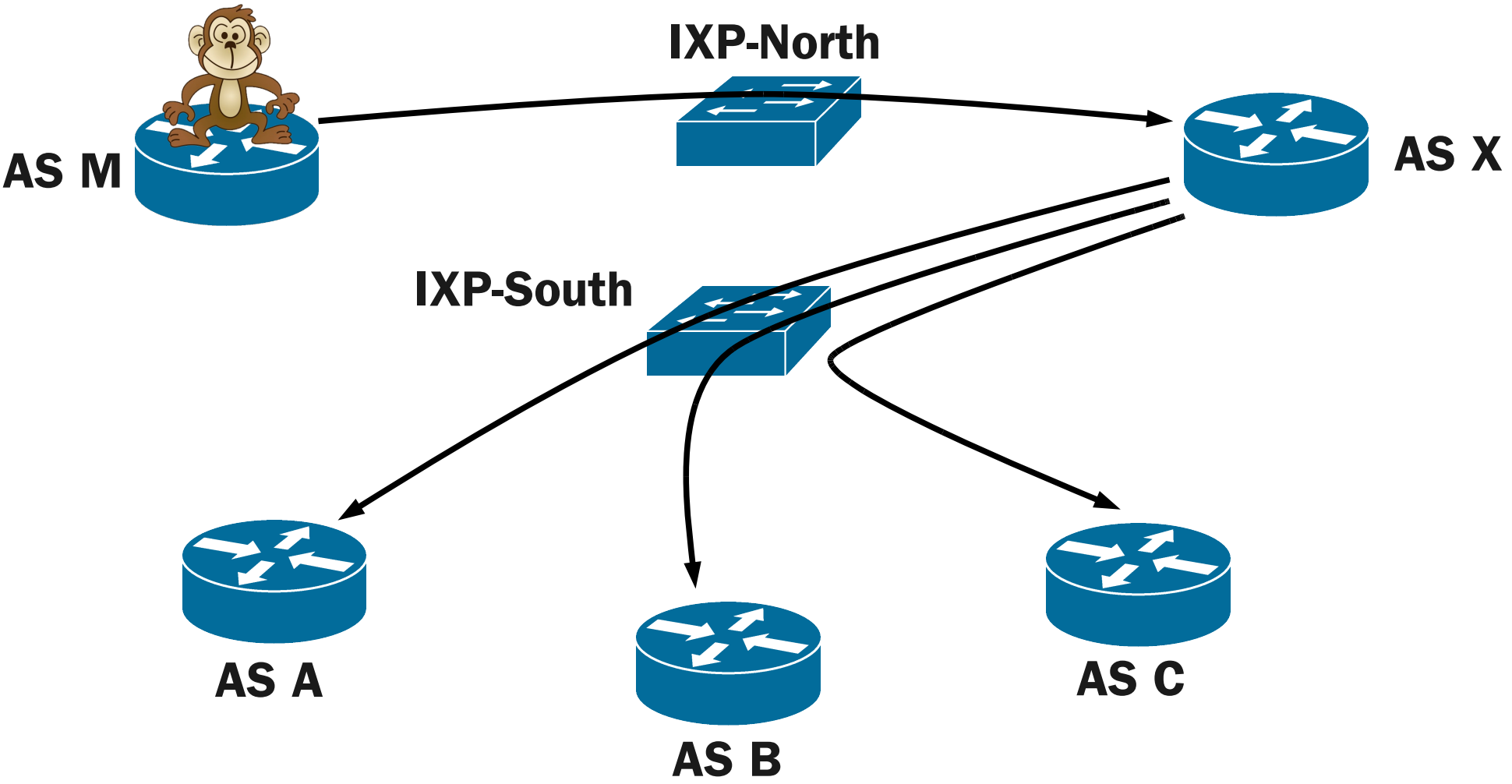
Borrowing other networks

- Into AS X to reach AS Y that is not present at the same IXP as AS M
  - utilise AS X's private peering link with AS Y



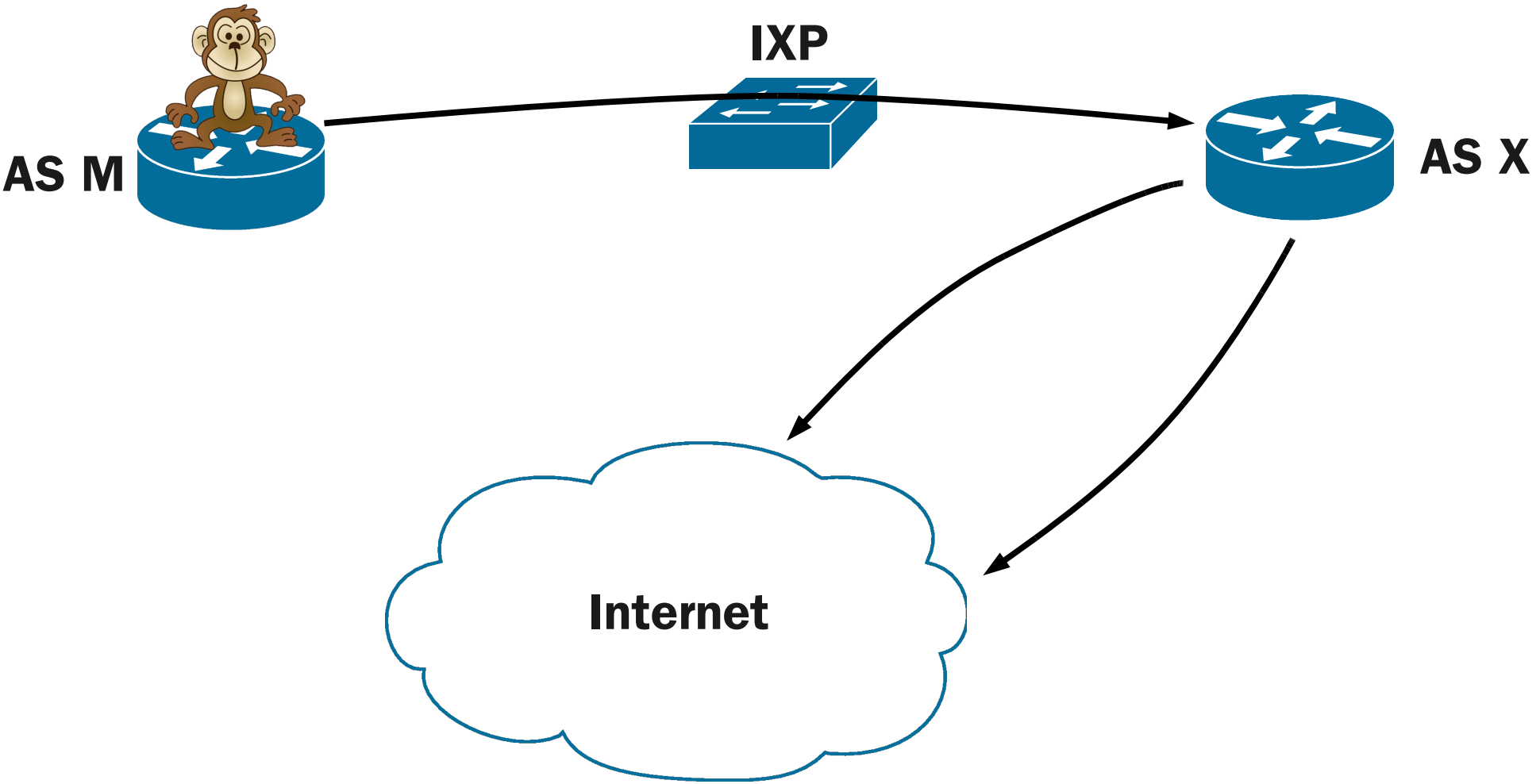
Borrowing other networks

- Into AS X to reach ASes at a different IXP



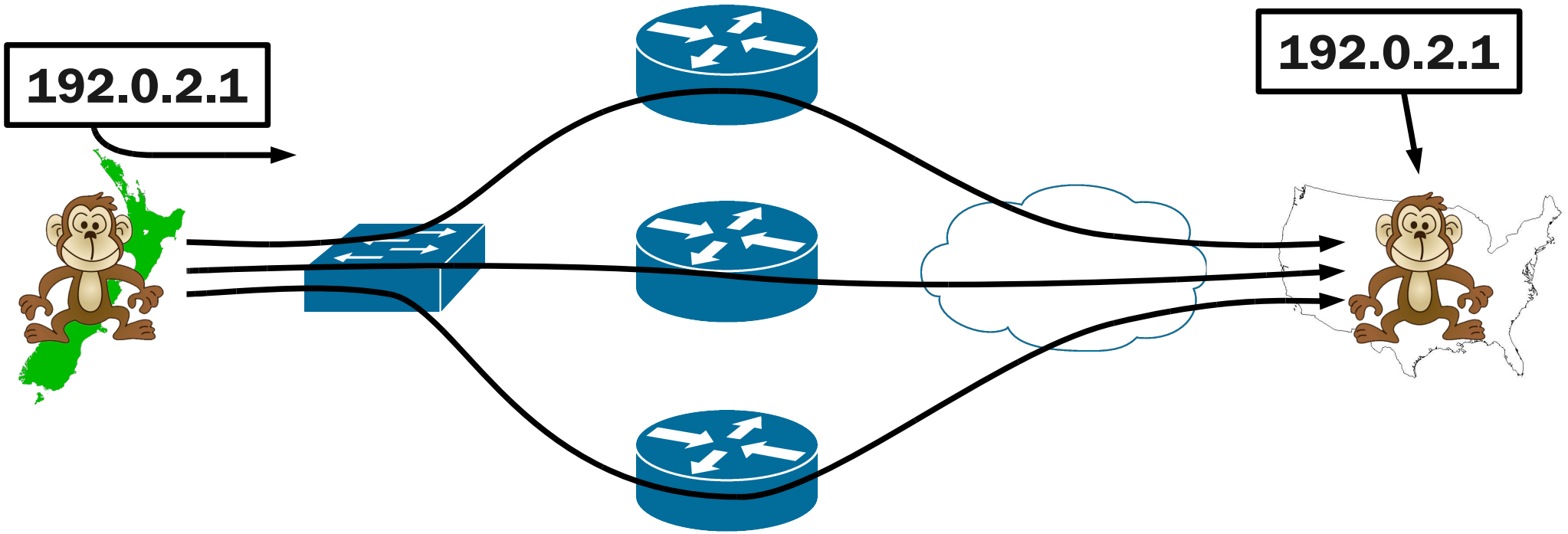
Borrowing other networks

- Into AS X to reach the internet at large



Borrowing other networks

- How many networks will deliver my packets?
  - "generally accepted" that this impolite, but how many networks actually protect themselves from it?



Borrowing other networks

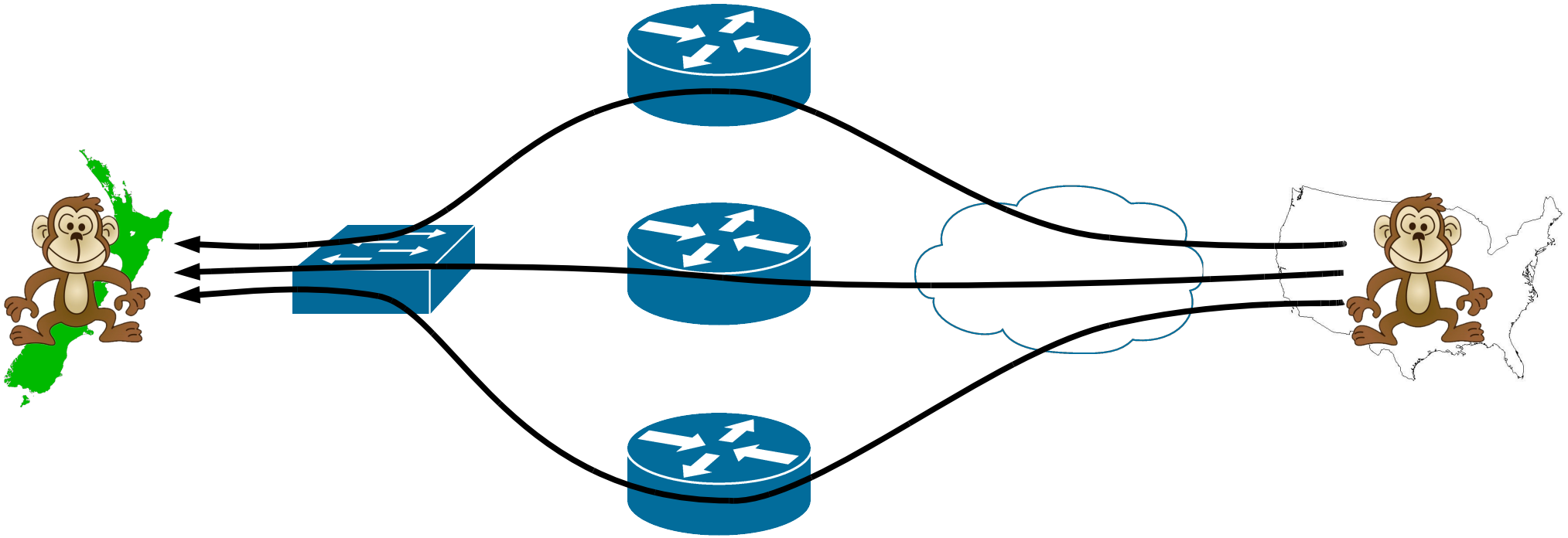
- How many networks will deliver my packets?
  - "generally accepted" that this impolite, but how many networks actually protect themselves from it?
- 78 hosts respond to ARP scan for APE /24
- 46 of those hosts carry packets to an international destination for me for free (59%!)
  - international transit in NZ costs ~US$100 - US$500 per Mbps depending on quantity purchased
  - oops.

Getting packets back

- So far, AS M can get packets out of their network to various destinations
- Great if AS M is a content provider, what if they are an access provider?
- Need to get packets back into their network
- A bit more complicated, can't just throw frames at another AS and hope for the best

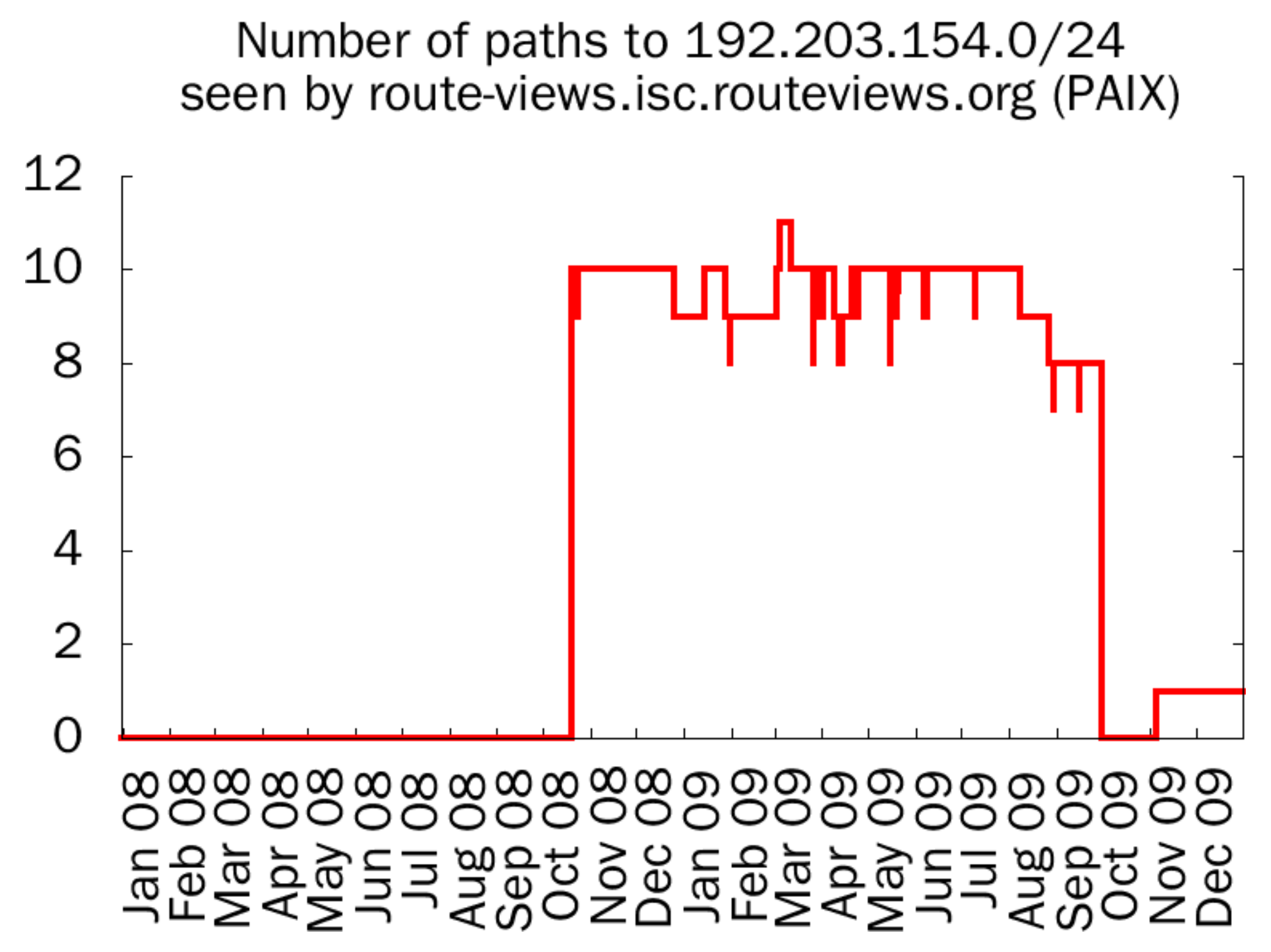
Getting packets back

- Requires that the far end has a path back that will pop out at the IXP



Getting packets back

- Packets sourced from IXP address requires far end to have a route to the IXP prefix
- Announcing an IXP prefix across an AS boundary is generally not a good idea
  - if you announce it upstream...
    - and your upstreams accept it
    - and they announce it to other networks/their upstreams
    - who announce it to yet more networks
    - etc
  - ...you're providing free transit for the IXP prefix


Number of paths to 192.203.154.0/24
seen by route-views.isc.routeviews.org (PAIX)
12
10
8
6
4
2
0
Jan 08
Feb 08
Mar 08
Apr 08
May 08
Jun 08
Jul 08
Aug 08
Sep 08
Oct 08
Nov 08
Dec 08
Jan 09
Feb 09
Mar 09
Apr 09
May 09
Jun 09
Jul 09
Aug 09
Sep 09
Oct 09
Nov 09
Dec 09

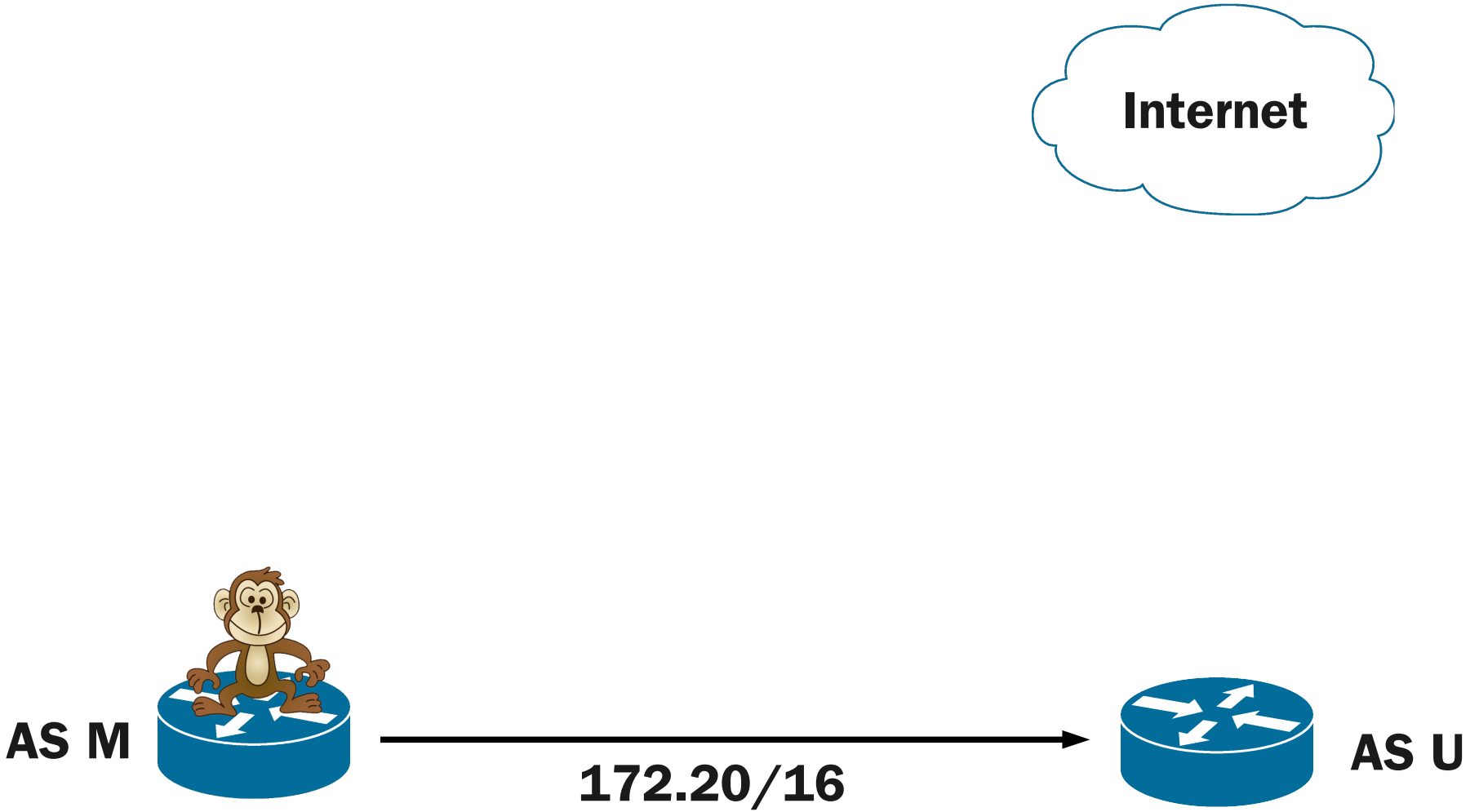
```
rs1.ape.citylink.co.nz-bgp> show ip bgp 192.203.154.0/24
BGP routing table entry for 192.203.154.0/24
Paths: (2 available, best #2, table Default-IP-Routing-Ta
 Advertised to non peer-group peers:
 *snip*
 24324 24324
  192.203.154.21 from 192.203.154.21 (202.147.41.170)
   Origin IGP, metric 0, localpref 100, valid, external
   Last update: Mon Feb 15 10:59:26 2010

 24324
  192.203.154.20 from 192.203.154.20 (203.192.167.174)
   Origin IGP, metric 0, localpref 100, valid, external,
   Last update: Mon Feb 15 10:59:23 2010
```

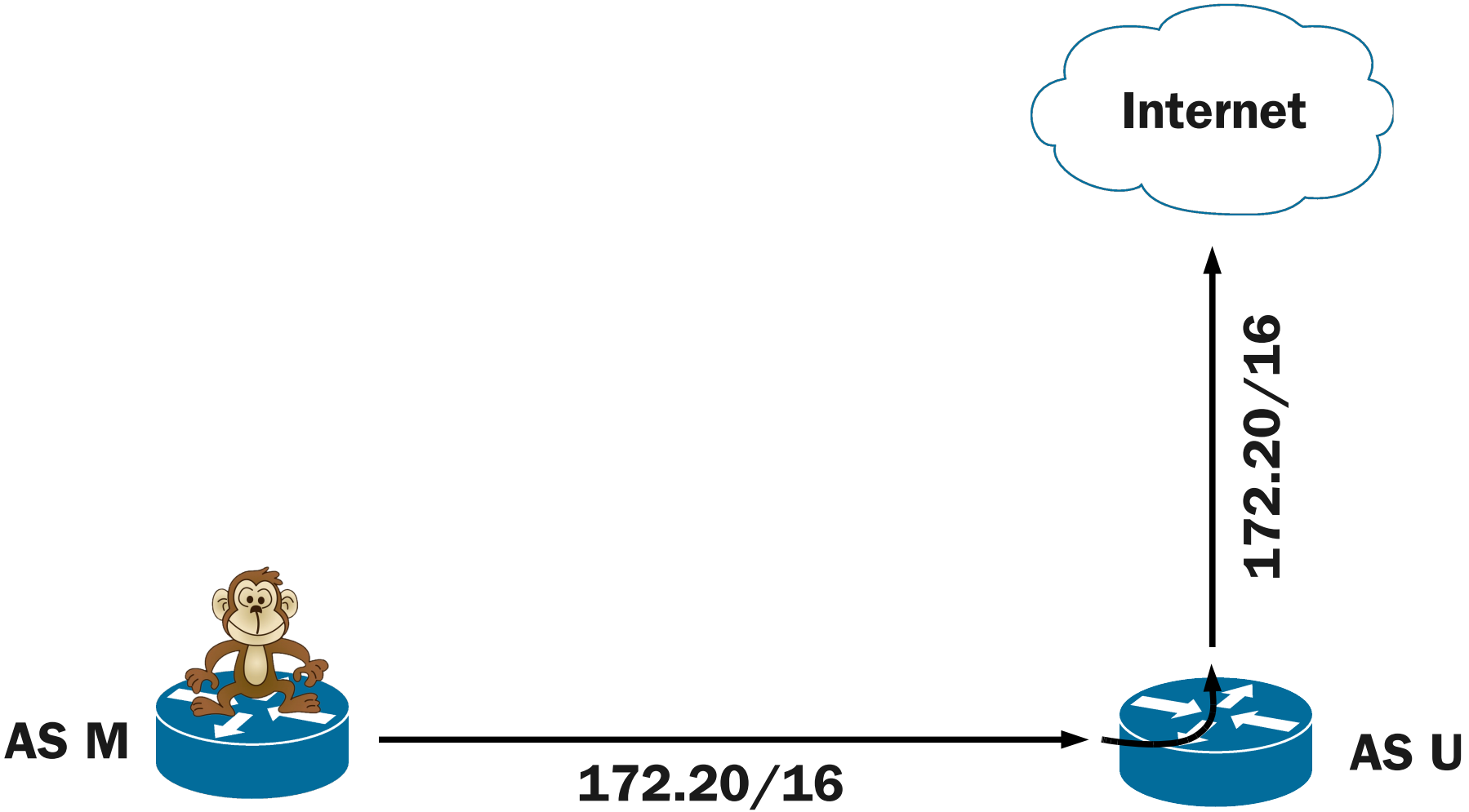
Getting packets back

- Packets sourced from IXP address requires far end to have a route to the IXP prefix
  - don't announce the IXP prefix outside your AS, really
- Packets sourced from non-IXP address requires far end to have a route to that address which will reach it via the IXP
  - (remember that router speaking OSPF at the IXP?)
  - play games with upstream provider's route table

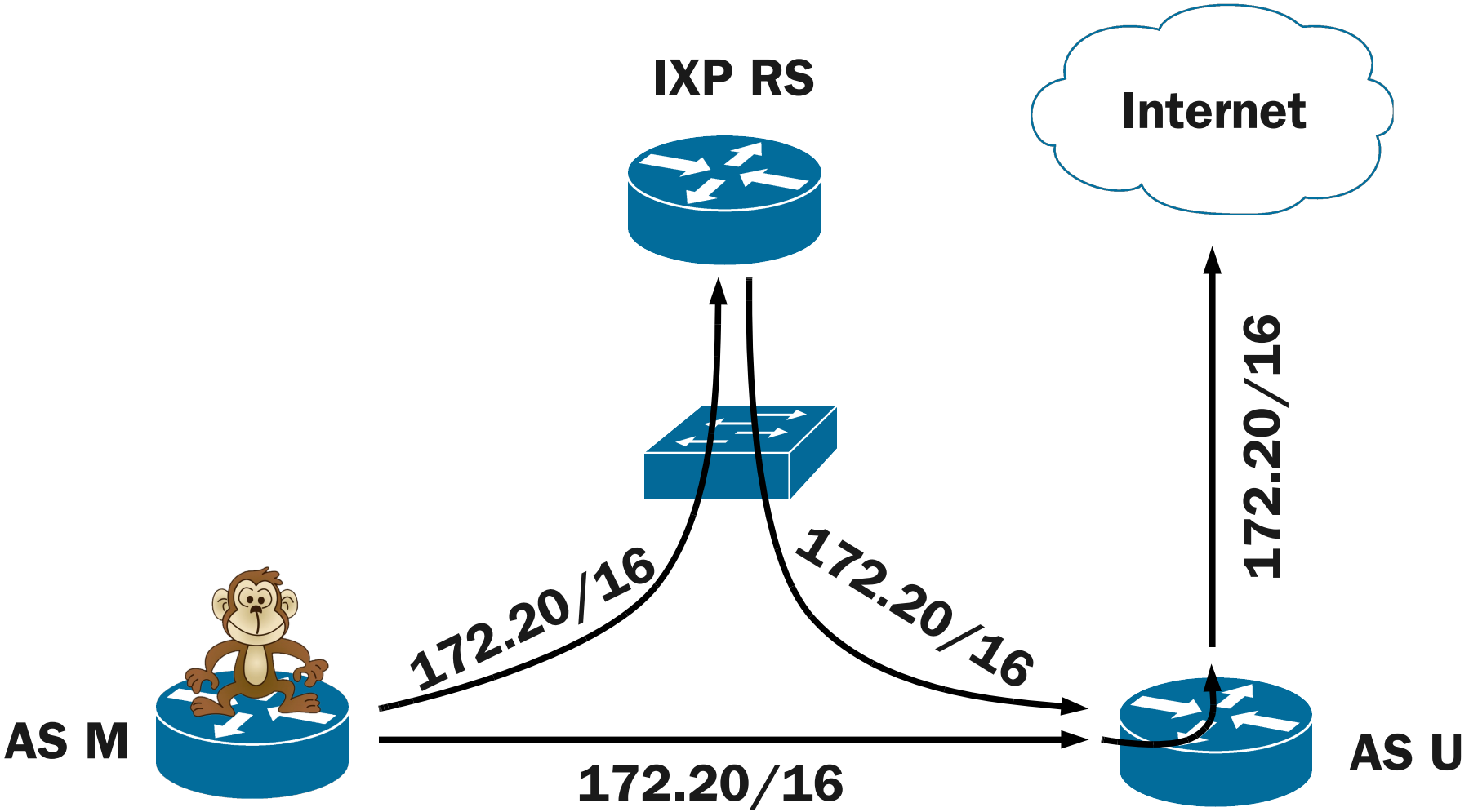
Getting packets back



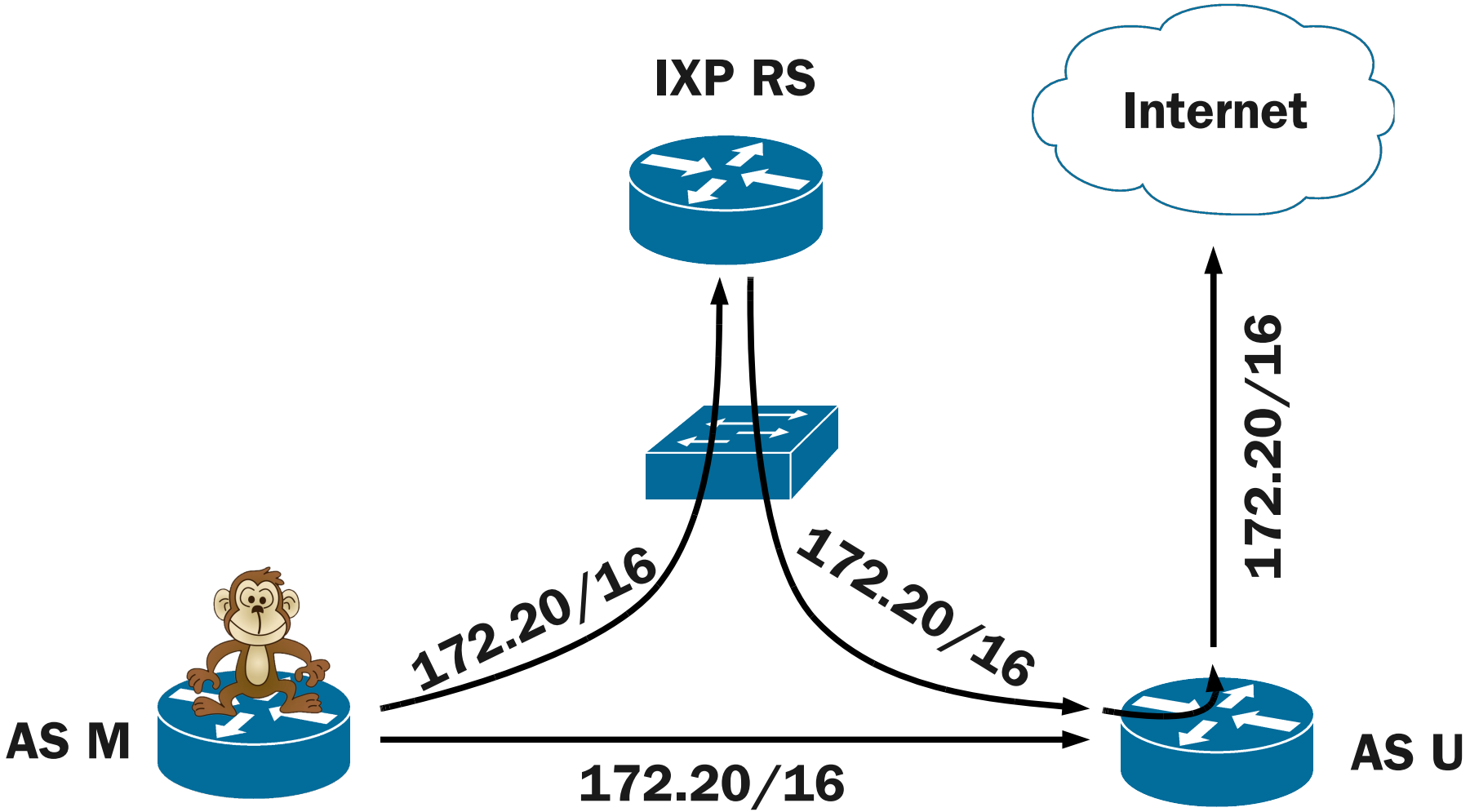
Getting packets back



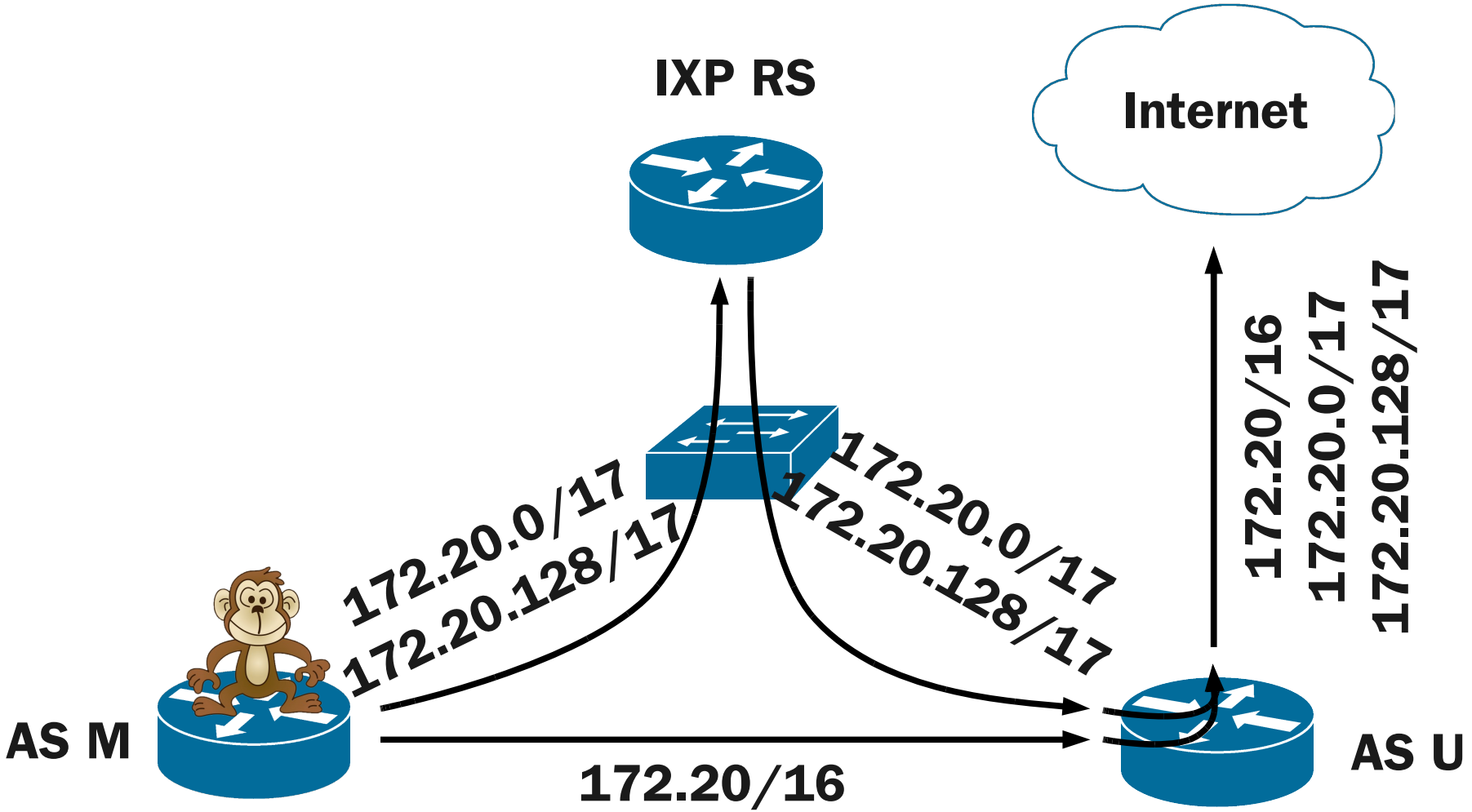
Getting packets back



Getting packets back

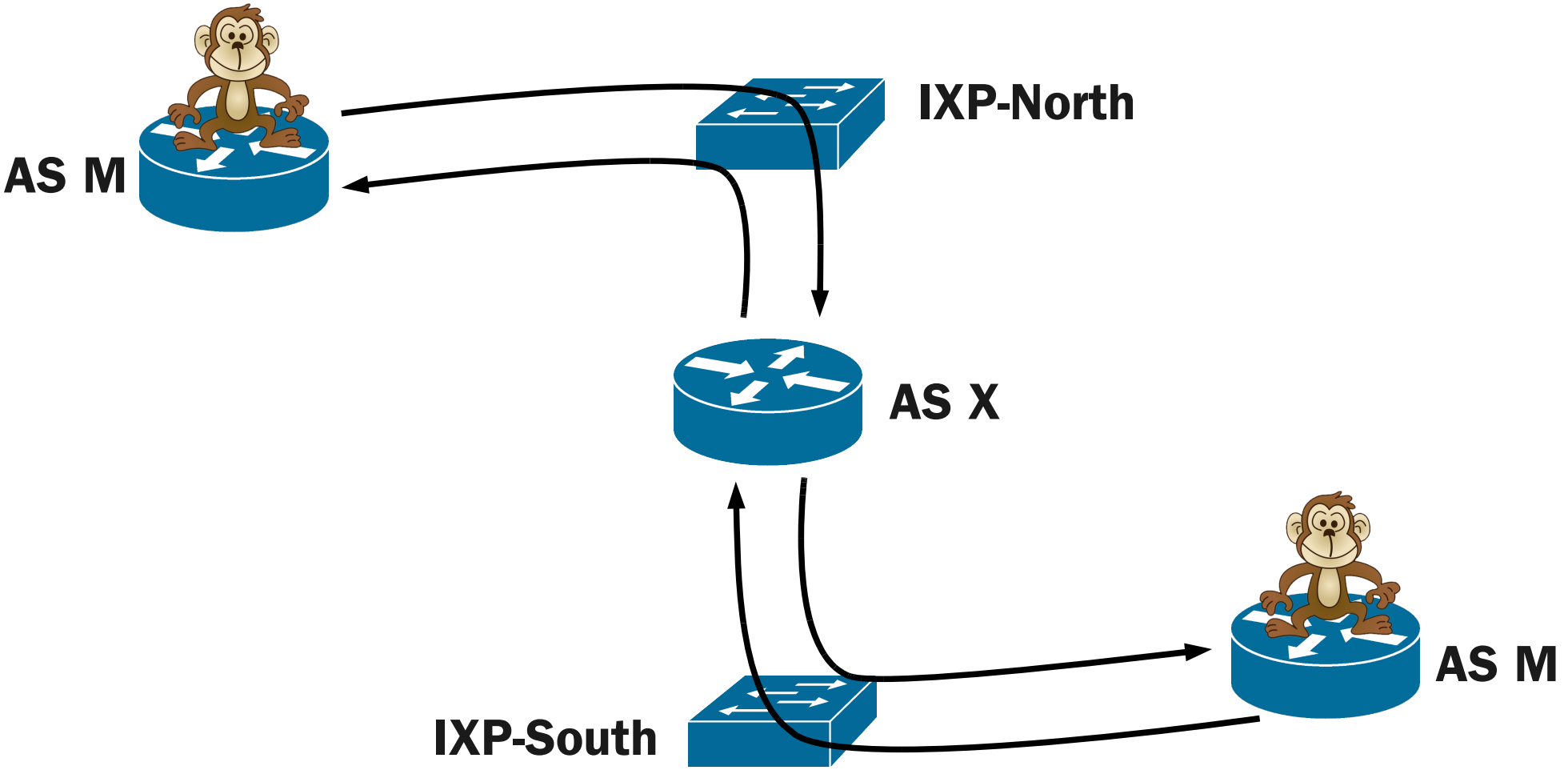


Getting packets back



Bidirectional flow

- Into AS X to reach yourself at a different IXP



New Zealand

~4.3 million people

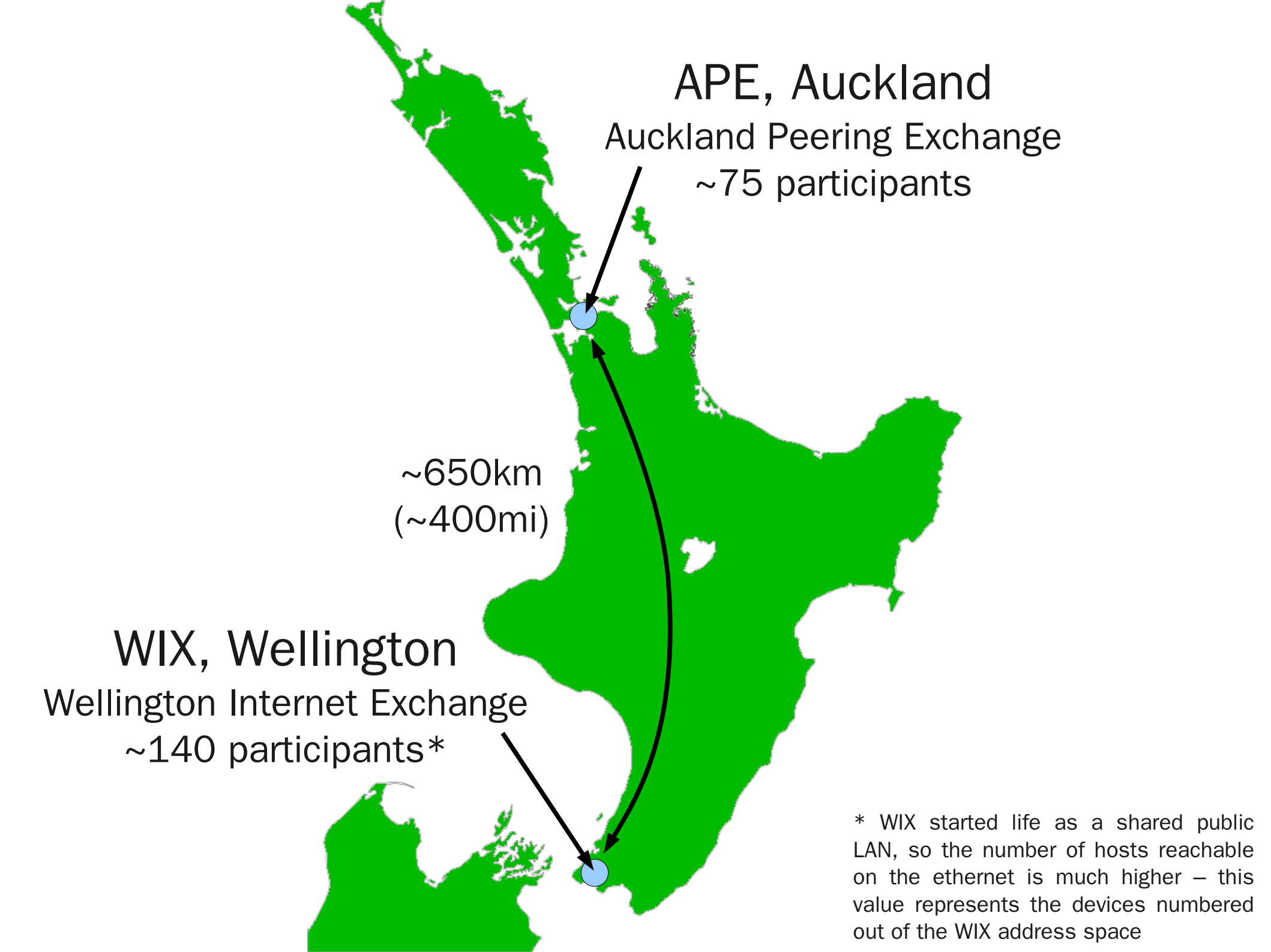
~43 million sheep

7 IXPs:

- 3CIX
- APE
- CHIX
- DPE
- HIX
- PNIX
- WIX

CITYLINK

www.citylink.co.nz



* WIX started life as a shared public LAN, so the number of hosts reachable on the ethernet is much higher – this value represents the devices numbered out of the WIX address space

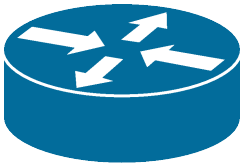
Borrowing other networks

- APE/WIX ports
  - ~US$100/month each for 1000 Mbps
- AKL – WLG ethernet service
  - ~US$800/month for 10 Mbps
- Almost an order of magnitude cheaper at 10 Mbps for someone to borrow your network than to extend their own!
  - even cheaper at 100 Mbps

Strong/Weak host model

- RFC1122, section 3.3.4.2.
  - "Proxy-ARP lite”

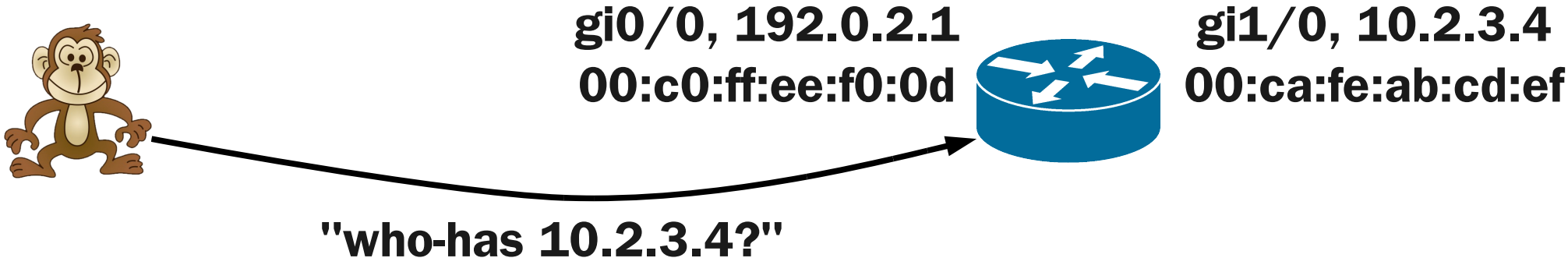
gi0/0, 192.0.2.1
00:c0:ff:ee:f0:0d



gi1/0, 10.2.3.4
00:ca:fe:ab:cd:ef

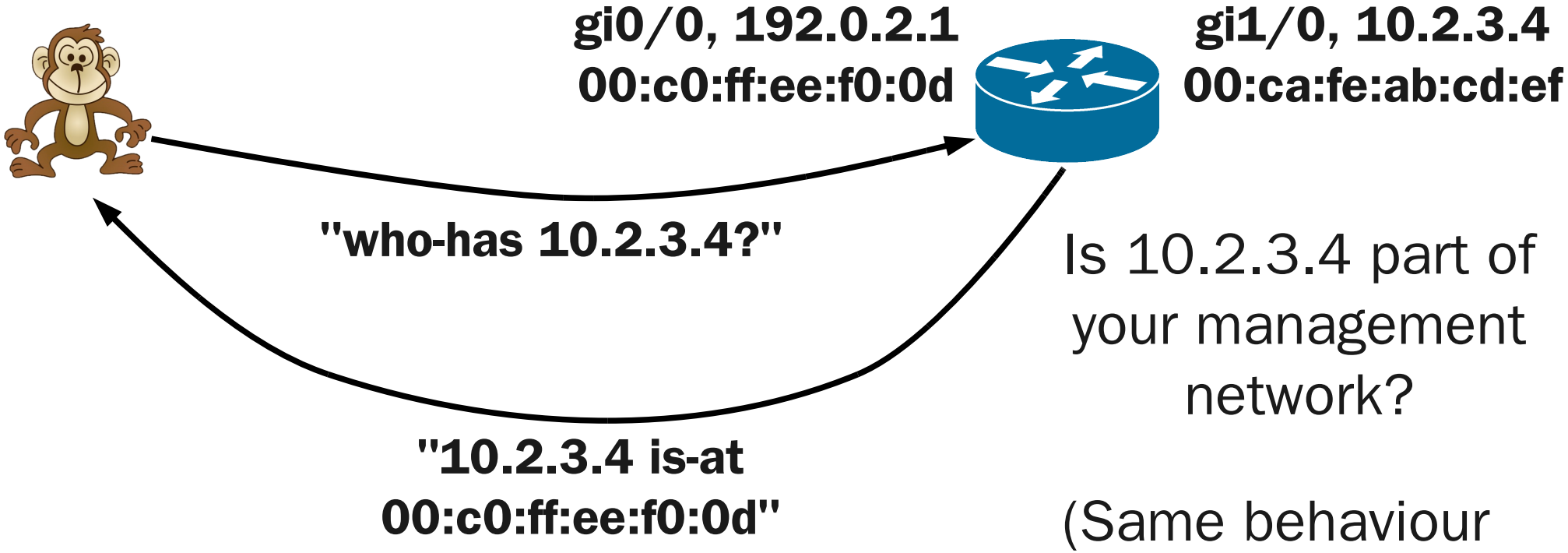
Strong/Weak host model

- RFC1122, section 3.3.4.2.
  - "Proxy-ARP lite”



Strong/Weak host model

- RFC1122, section 3.3.4.2.
  - "Proxy-ARP lite”



Causing trouble

- responding to DHCP requests
  - "Today you will be 10.20.30.40, and you should use me as your default gateway"
- sending IPv6 RAs, or answering IPv6 RSs
  - roughly same effect as DHCP, but for v6
- responding to un-answered ARP requests
  - stand up a BGP session with someone who has left an old neighbour configured in their router

Going to jail

- ARP spoofing
  - "Hi, I'll be your new route servers today"
  - (severe lack of MD5-enabled BGP sessions)
- speaking OSPF back to the OSPF-speaker
  - "I can reach 8.8.8.8/32, send me packets"
  - are they redistributing OSPF into BGP?
- sending TCP RSTs
  - RFC3682 (TTL hack) wont save you here

How do we fix it?

- Go back to last week, and attend one of the routing workshops
- Get the workshop slides, read them
- AMS-IX configuration guide

http://ams-ix.net/config-guide

How do we fix it?

- Make sure you're only accepting packets you want to be accepting
  - probably only want packets destined for networks that you're advertising via BGP
  - in most cases, only want packets destined for your network, and networks you sell transit to
    - selling transit via the IXP is asking for trouble
    - so is advertising the IXP prefix outside of your network

How do we fix it?

- Stop packets you don't want using your network
- Method depends on what your network does, how complicated it is, and your budget
- Either:
  - prevent unwanted packets entering your network
    - apply filters on your peering router(s) to only allow packets destined for "valid" destinations through
  - ensure unwated packets can't get anywhere useful
    - ensure that your peering router(s) only have routes for "valid" destinations in its/their route table

How do we fix it?

- "Stub" network
  - smaller networks, only one or two routers in total
  - not many IP customers (and not much churn)
  - unlikely to have dedicated peering router
- Router probably carries default route
- Apply filters on IXP interface so only packets destined for your network (and customers) are allowed in

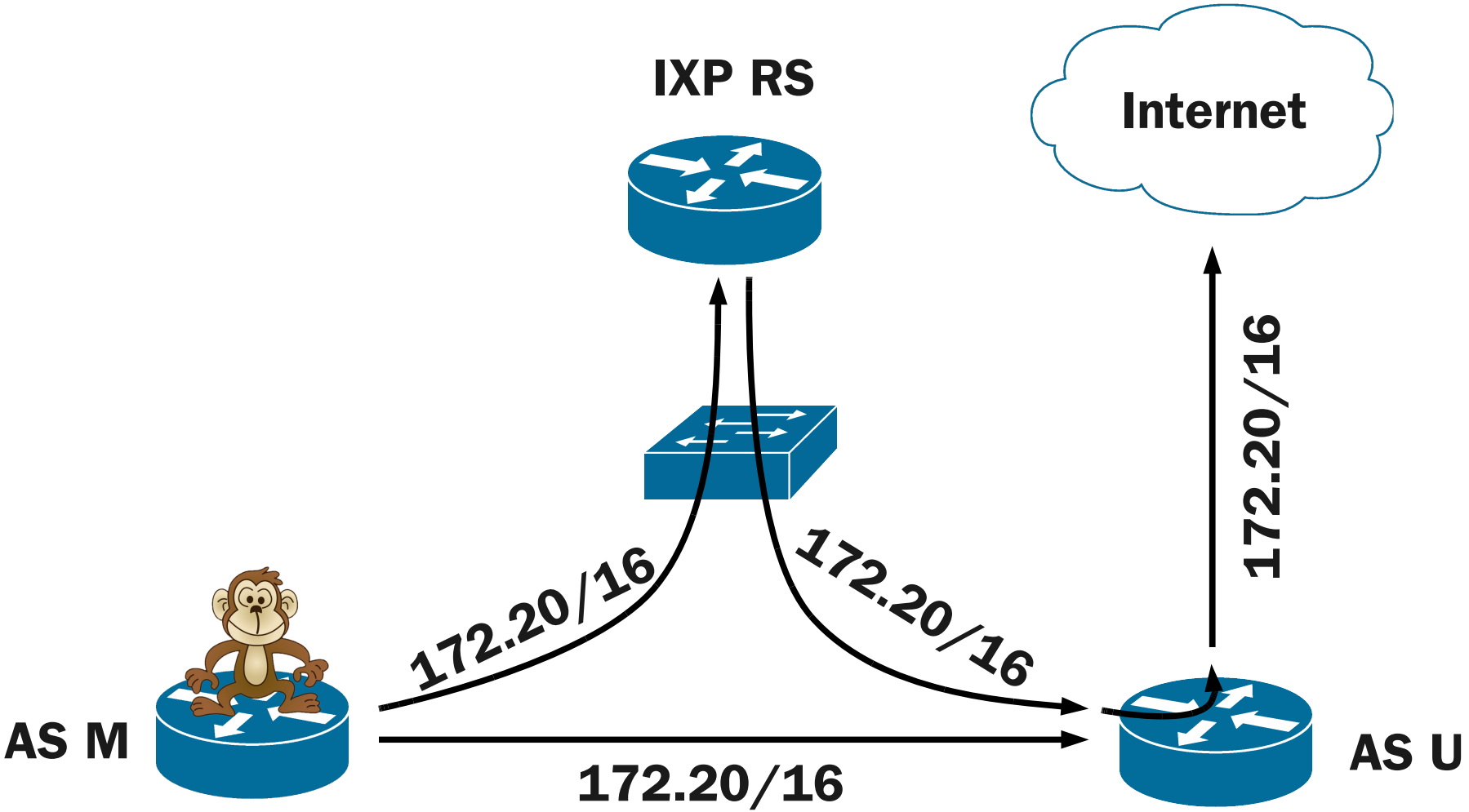
How do we fix it?

- Network with dedicated peering router
  - modifying IXP interface packet filters may be too much work as customers come and go
    - but other things need modification when customers come and go anyway, so automate it as part of your provisioning and deprovisioning processes
- Ensure router carries only:
  - your prefixes
  - prefixes learned from peers at IXP
- Router must have no default route

How do we fix it?

- Large network, complex routing policies
- Multiple routing tables
- Ensure that your IXP interface is in a VRF that only contains:
  - your prefixes
  - prefixes learned from peers at IXP

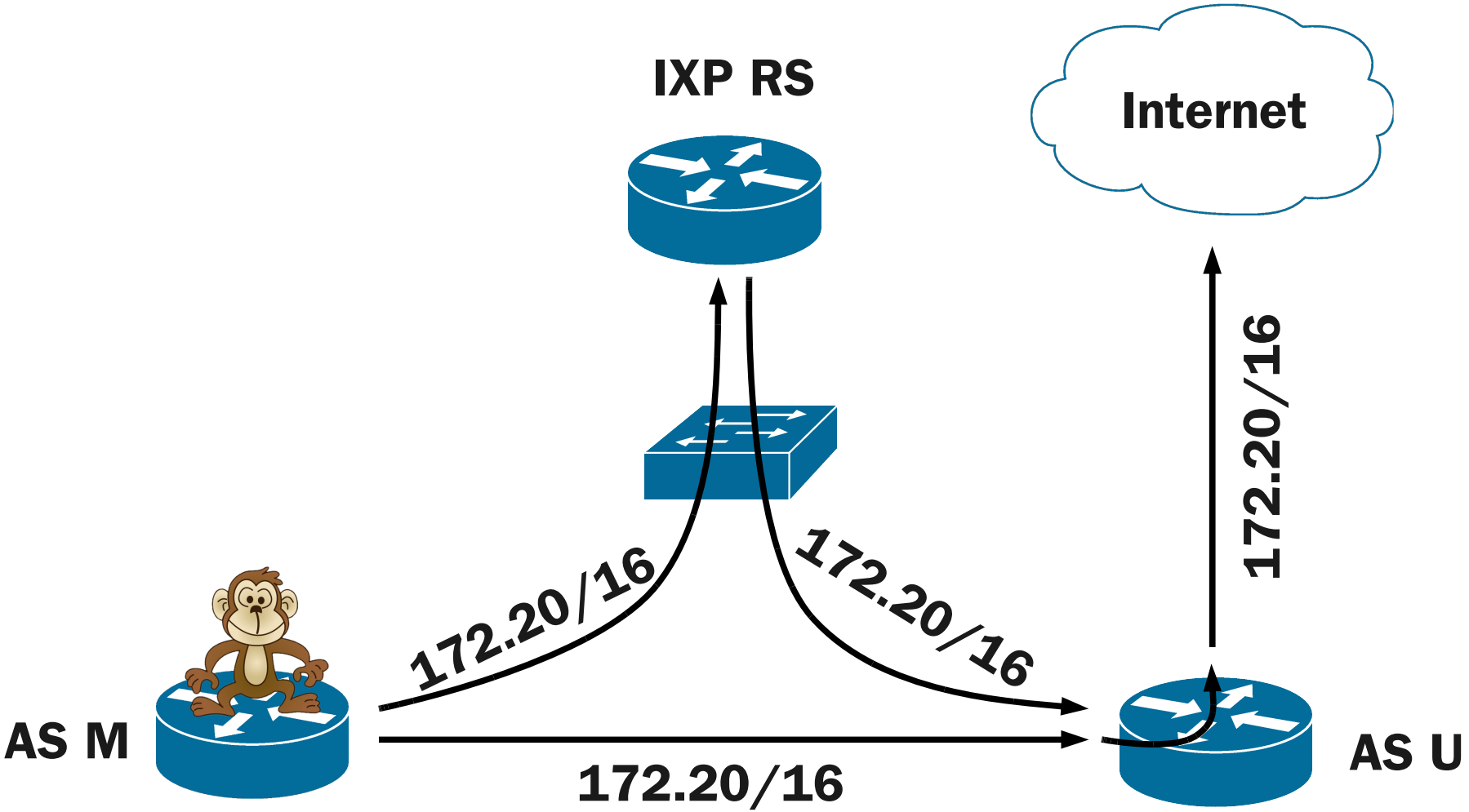
How do we fix it?



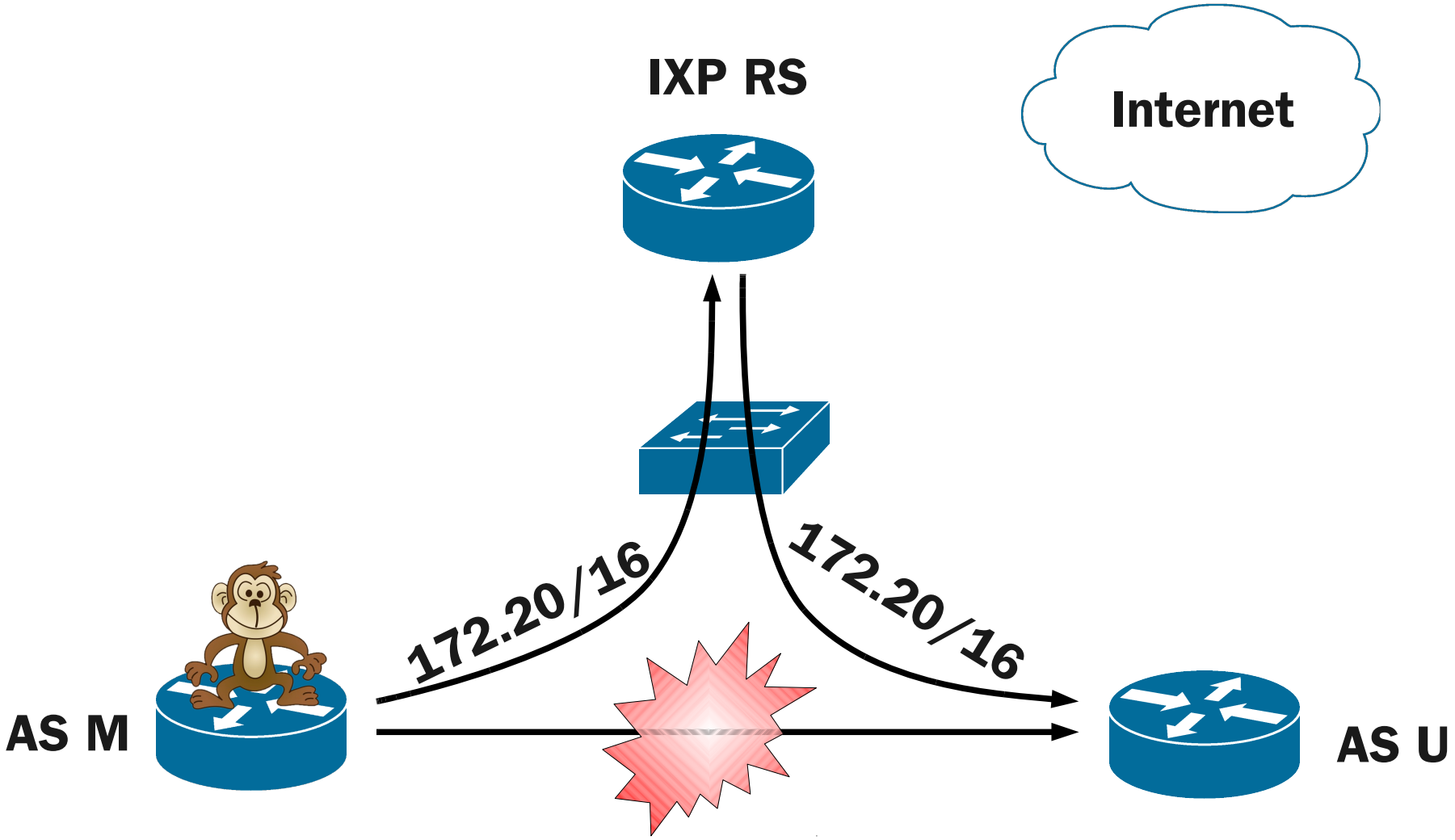
How do we fix it?

- Generally you don't want to use any prefixes received via an IXP to reach a customer that you're also providing transit to
- However, if you lose your link to your customer, you may now wish to reach them via the IXP
  - but you probably still don't want to provide transit to them over the IXP
- Make use of VRFs - you can reach the customer via the IXP for packets from your network to theirs, and via your link to them for transit

How do we fix it?



How do we fix it?



When connecting your network to a shared layer 2 network such as an Internet Exchange Point, make sure you take appropriate steps to protect your network against misuse

(You want to take steps to protect your network from misuse even when it's not attached to an IXP, but there are more ways people can fiddle with your network when, for a small monthly fee, they can get on the same layer 2 network as one of your routers)

Questions?

Mike Jager
mike@mikej.net.nz