BGP Best Practices for ISPs

Prefix List, AS PATH filters, Bogon Filters, Anycast, Mailing Lists, INOC DBA

....

Gaurab Raj Upadhaya

gaurab@pch.net

Packet Clearing House

PCH
Packet Clearing House

What are Best Practices

- Established or known good ways of doing things.
- Being a good Internet citizen.
- Trying TO restrict damage to your network and FROM you network
- A lot of RFCs devoted on different BCP, but here we'll just cover a few topics here

Packet Clearing House

BGP Security

Packet Clearing House

Protocols are vulnerable

- Routing protocols, like BGP, OSPF can be attacked and weakness exploited
  - False Information
  - Man-in-middle situation
  - Denial of service
  - Routing Conditions

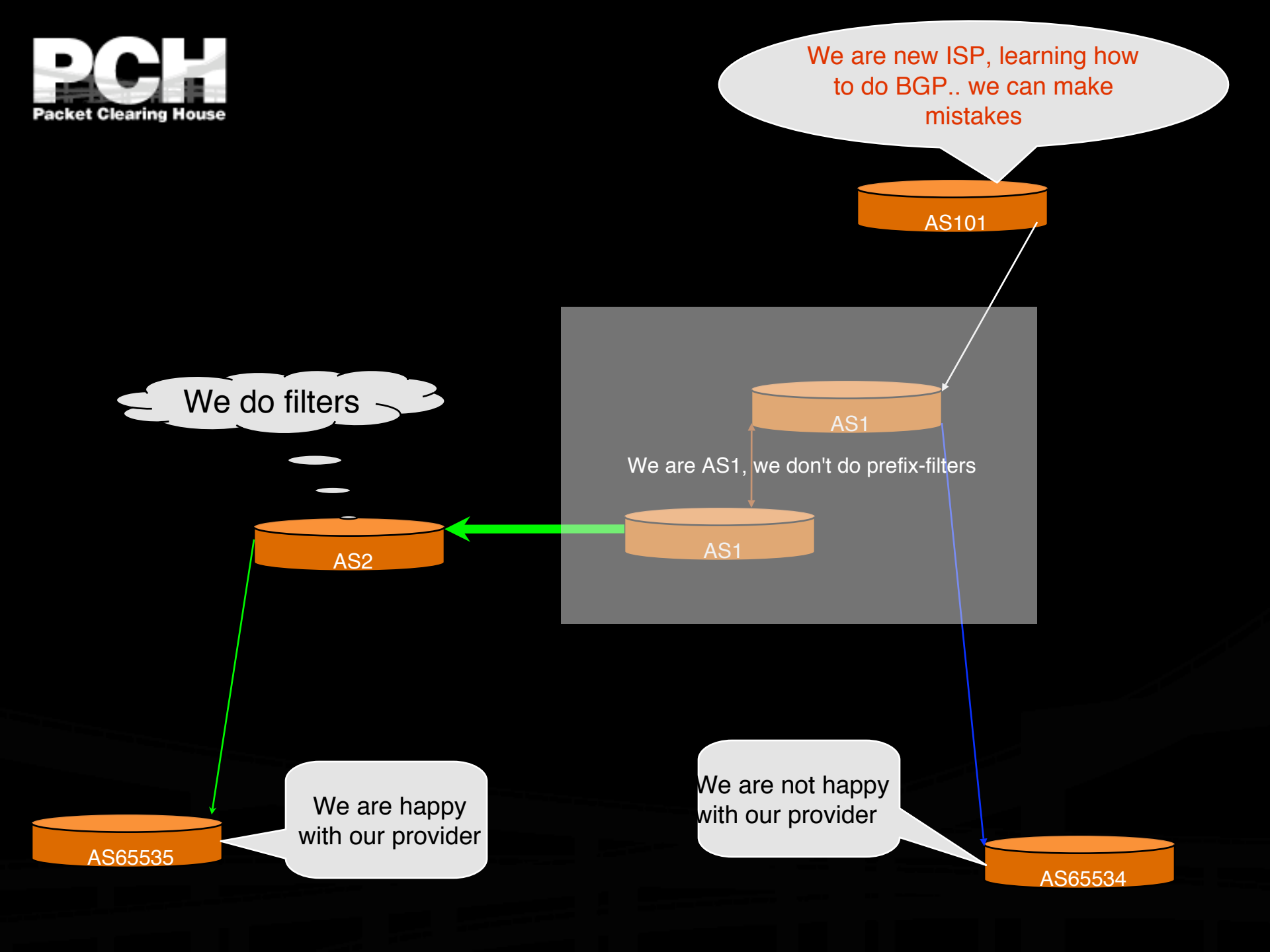
Packet Clearing House

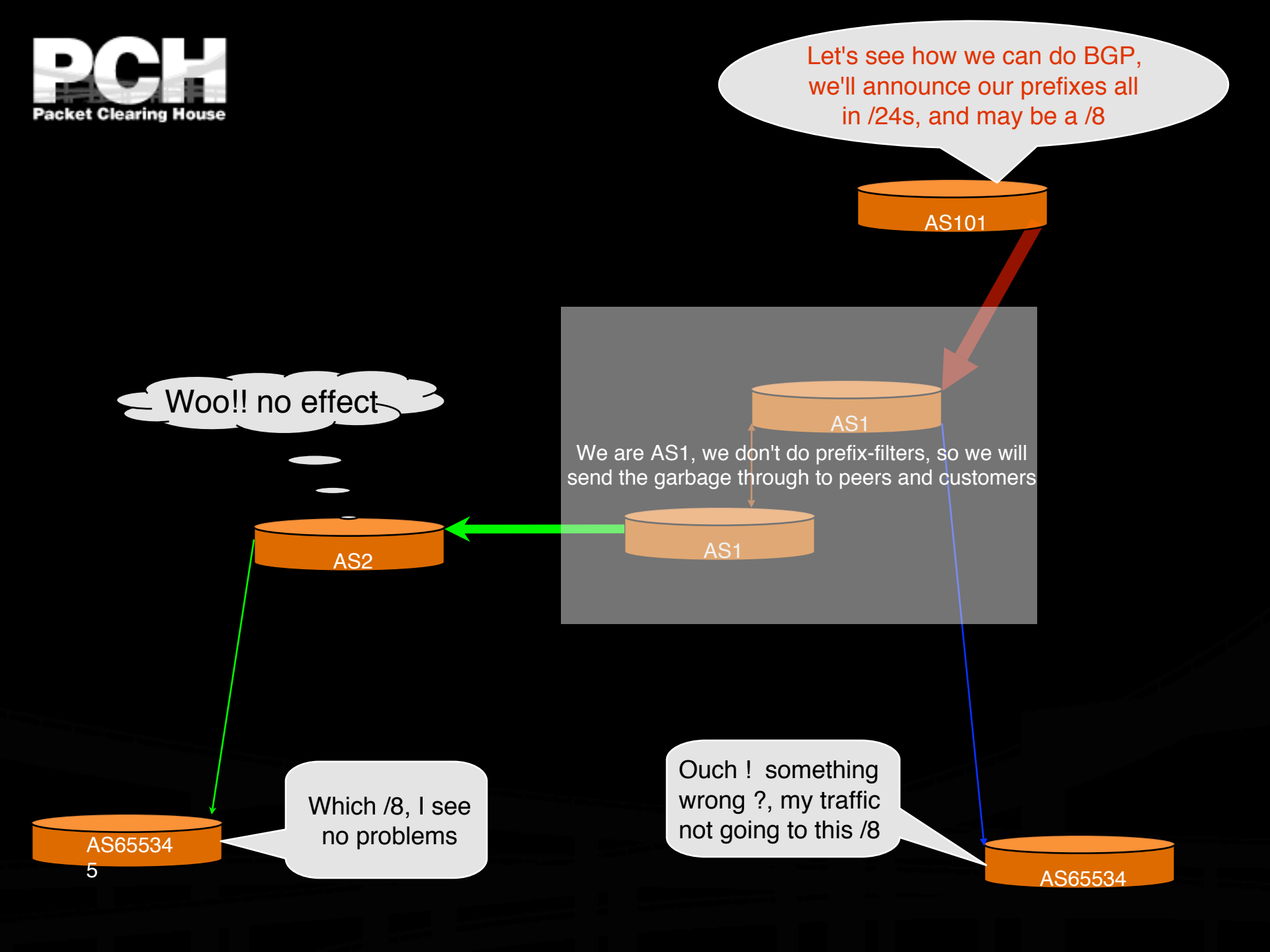
BGP Security Issues

- Prefix Lists
  - Why, How and Where
- Prefix Lists
  - Customers
  - Peers
  - Upstream
- Protocol Authentication
  - MD5
- BGP BCP
  - Bogons

Packet Clearing House

Why Prefix Filter?


PCH
Packet Clearing House
We are new ISP, learning how to do BGP.. we can make mistakes
AS101
We do filters
AS1
We are AS1, we don't do prefix-filters
AS1
AS2
AS65535
We are happy with our provider
We are not happy with our provider
AS65534


Packet Clearing House
Let's see how we can do BGP, we'll announce our prefixes all in /24s, and may be a /8
AS101
AS1
We are AS1, we don't do prefix-filters, so we will send the garbage through to peers and customers
AS1
Woo!! no effect
AS2
AS65534
Which /8, I see no problems
Ouch ! something wrong ?, my traffic not going to this /8
AS65534

PCH
Packet Clearing House

It's REAL, does happen

- In 1997, this happened with AS 7007 - which was the most visible of all.
- Frequent Messages on IXes mailing list about hitting max prefix limits.
- Garbage in – Garbage out.

PCH
Packet Clearing House

So, what should we do ?

- Don't send Junk
  - Filter your own advertisement
  - Filter your customers
- Don't Accept Junk
  - Filter your customers
  - Filter your peers
- Help Others mitigate impact - Net Police
  - Help others and they'll help you

Packet Clearing House

Prefix Filter - How

BGP Prefix Filtering

- Two ways to implement prefix filtering:
  - Prefix list - easy to use, not highly scalable
  - AS-PATH - widely used, highly scalable
    - Distribute list is now considered obsolete
- Two filtering techniques:
  - Explicit Permit (permit then deny any)
  - Explicit Deny (deny then permit any)

Packet Clearing House

BGP Route Filtering

Prefix-List example

```
ip prefix-list rfc1918 deny   0.0.0.0/8 le 32
ip prefix-list rfc1918 deny   10.0.0.0/8 le 32
ip prefix-list rfc1918 deny   127.0.0.0/8 le 32
ip prefix-list rfc1918 deny   169.254.0.0/16 le 32
ip prefix-list rfc1918 deny   172.16.0.0/12 le 32
ip prefix-list rfc1918 deny   192.0.2.0.0/24 le 32
ip prefix-list rfc1918 deny   192.168.0.0/16 le 32
ip prefix-list rfc1918 deny   224.0.0.0/3 le 32
ip prefix-list rfc1918 permit 0.0.0.0/0 le 32
```

Packet Clearing House

BGP Filtering

Applying the previous prefix-list to peers

```
router bgp 65564
 no synchronization
 neighbor 198.32.231.200 remote-as 65200
 neighbor 198.32.231.200 prefix-list rfc1918 in
 neighbor 198.32.231.200 prefix-list rfc1918 out
 neighbor 198.32.231.210 remote-as 65210
 neighbor 198.32.231.210 prefix-list rfc1918 in
 neighbor 198.32.231.210 prefix-list rfc1918 out
 no auto-summary
!
```

Packet Clearing House

Using AS-PATH filters

Using AS-PATH filters

- Filter routes based on AS path
  - Applied same way as prefix-list filters
  - AS-PATH syntax is different, can use regular expressions
- Example Configuration:

```
router bgp 65564
 network 10.12.0.0 mask 255.255.0.0
 neighbor 198.32.231.200 remote-as 65200
 neighbor 198.32.231.200 filter-list 1 out
 neighbor 198.32.231.200 filter-list 200 in
!
ip as-path access-list 1 permit ^65564$
ip as-path access-list 200 permit ^65200$
```

Regular Expressions

- Most router OS uses Unix regular expressions

.	Match one character
*	Match any number of preceding expression
+	Match at least one of preceding expression
^	Beginning of line
$	End of line
_	Beginning, end, white-space, brace
|	Or
( )	brackets to contain expression

Regular Expressions examples

- Examples

.*	match anything
.+	match at least one character
^$	match routes local to this AS
_3856$	originated by AS3856
^3856_	received from AS3856
_3856_	via AS3856
_3856_42_	via AS3856 and AS42
_(3856_)+	multiple AS3856 in sequence (used to match AS-PATH prepends)

Regular Expressions examples

- Complex Examples

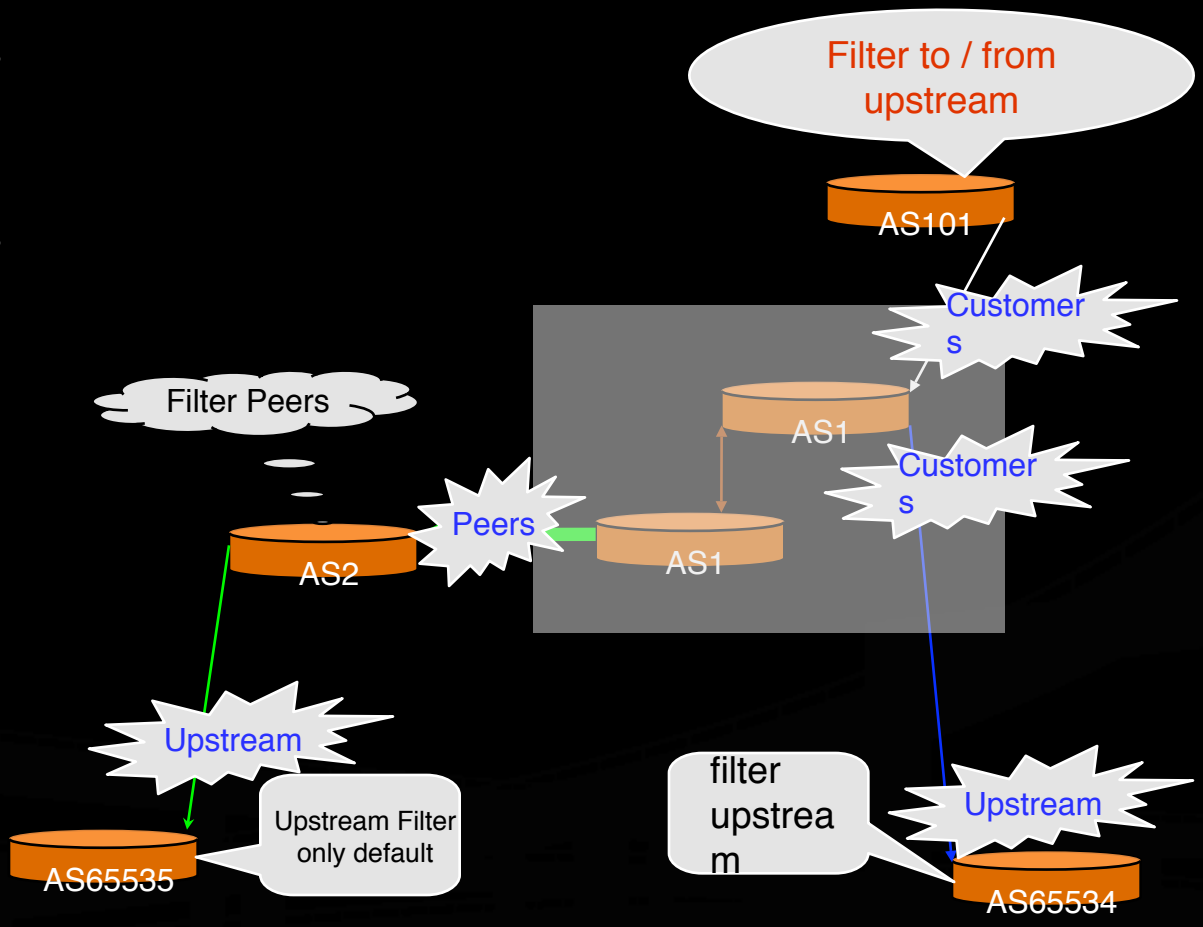
^[0-9]+$	Match AS_PATH length of one
^[0-9]+_[0-9]+$	Match AS_PATH length of two
^[0-9]*_[0-9]+$	Match AS_PATH length of one or two
^[0-9]*_[0-9]*$	Match AS_PATH length of one or two (will also match zero)
^[0-9]+_[0-9]+_[0-9]+$	Match AS_PATH length of three
_(3856|42)_	Match anything which has gone through AS42 or AS3856
_2914(_.+_)42$	Match anything of origin AS42 and passed through AS2914

PCH
Packet Clearing House

Where to Prefix Filter?

Where to Prefix Filter?

- Customers
  - Ingress / Egress
- Upstream
  - Ingress / Egress
  - Use default if single homed
- Peers
  - Ingress and Egress at all points of peering



PCH
Packet Clearing House

Special Use Addresses

PCH
Packet Clearing House

Special Use Addresses

- There are routes that should NOT be routed on the Internet
  - RFC 1918 and “Martian” networks
  - 127.0.0.0/8 and multicast blocks
  - Certain RFC3330 addresses:
  - http://www.rfc-editor.org/rfc/rfc3330.txt
- BGP should have filters applied so that these routes are not advertised to or propagated through the Internet

PCH
Packet Clearing House

Special Use Addresses

- Quick review
  - 0.0.0.0/8 and 0.0.0.0/32—Default and broadcast
  - 127.0.0.0/8—Host loopback
  - 192.0.2.0/24—TEST-NET for documentation
  - 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16—RFC 1918 private addresses
  - 169.254.0.0/16—End node auto-configuration

Packet Clearing House

Special Use Addresses

```
ip prefix-list deny-sua deny 0.0.0.0/8 le 32
ip prefix-list deny-sua deny 10.0.0.0/8 le 32
ip prefix-list deny-sua deny 127.0.0.0/8 le 32
ip prefix-list deny-sua deny 169.254.0.0/16 le 32
ip prefix-list deny-sua deny 172.16.0.0/12 le 32
ip prefix-list deny-sua deny 192.0.2.0/24 le 32
ip prefix-list deny-sua deny 192.168.0.0/16 le 32
ip prefix-list deny-sua deny 224.0.0.0/3 le 32
ip prefix-list deny-sua deny 0.0.0.0/0 ge 25
ip prefix-list deny-sua permit 0.0.0.0/0 le 32
```

PCH
Packet Clearing House

Bogons

- IANA publishes the list of IP Address that have been assigned to RIRs and end-users
  - http://www.iana.org/assignments/ipv4-address-space
- Only these blocks of IPv4 addresses should be visible and used on the Internet.
- Filters should be applied on Ingress / Egress of your AS for all other address range

PCH
Packet Clearing House

Bogons

- The IP Address is 32 Bits. The range is
  - 0.0.0.0 to 255.255.255.255
- Not all IP Address are allocated by IANA
- Few are not to be used on the public internet (RFC 1918)
- Few blocks are reserved (> 223/8)
- The IP Address that are not supposed to be seen on the Internet are known as ‘Bogons’, sometime also referred to as “Martian”

PCH
Packet Clearing House

Bogons

- To be a good Internet citizen, ISPs should not route them
  - Many DoS attacks originate from these unallocated address blocks, so it's also good security
- ISPs can get the bogons list from IANA and set up their prefix filters to route traffic
- When new address is assigned, generally the announcements are sent far and wide on many different mailing lists (eg. sanog, nanog, apops, afnog etc.)
- Same procedure is also applied for IPv6 address space

PCH
Packet Clearing House

Router Configuration

- This is how you configure your Cisco routers in your networks
  - Static Route to Null0, good idea for RFC1918 space
    - ip route 192.0.2.0 mask 255.255.255.0 null0
  - BGP prefix-list for Unallocated blocks
    - neighbor x.x.x.x prefix-list bogons
    - ip prefix-list bogons seq 5 deny 10.0.0.0/0 le 32
    - ip prefix-list bogons seq 10 deny 127.0.0.0/0 le 32

PCH
Packet Clearing House

Router Configuration

- The problem with the static configuration is many fold
  - The Bogons list keeps on changing
  - People move into new jobs, new people start managing the network
  - New engineers may have no clue on why those configurations are in there
  - The network is working, so let's leave it there triumphs over further digging

PCH
Packet Clearing House

Bogon Route Server Project

- Thus, the Bogon Route Server Project
  - Provides bogons over a eBGP Multihop session. There are four Bogons Route Server. Connecting to at least two is recommended.
  - You can run a BGP session with the Bogon route server
  - You receive the bogons list through BGP, then you can either
    - Either set the next-hop for the bogons to a static IP address and help in research work
    - Or set the next-hop for the bogons to a static IP address that is statically routed to Null0

Configuring Routers for Bogon RS

- Full Details are available at http://www.cymru.com/BGP/bogon-rs.html

```
router bgp <your asn>
   neighbor x.x.x.x remote-as 65333
   neighbor x.x.x.x ebgp-multihop 255
   neighbor x.x.x.x description <your description>
   neighbor x.x.x.x prefix-list cymru-out out
   neighbor x.x.x.x route-map CYMRUBOGONS in
   neighbor x.x.x.x password <your password>
   neighbor x.x.x.x maximum-prefix 100 threshold 90
!
ip bgp-community new-format
!
ip route 192.0.2.1 255.255.255.255 null0
!
ip community-list 10 permit 65333:888
!
route-map CYMRUBOGONS permit 10
   description Filter bogons learned from cymru.com bogon route-servers
   match community 10
```

Packet Clearing House

How do I set up the peering ?

- To peer with the Bogon Route Server,
  - contact team-cymru@cymru.com.
- When requesting a peering session, you should include the following information in your email:
  - Your AS number
  - The IP address(es) you want to use for peering
  - If your equipment support MD5 passwords for BGP sessions?
  - Your PGP/GPG public key, if you have one (not mandatory)
- The session is set up through eBGP multihop with a private ASN. It currently has 95 prefixes. Your router must at least support these requirements.

PCH
Packet Clearing House

Team Cymru

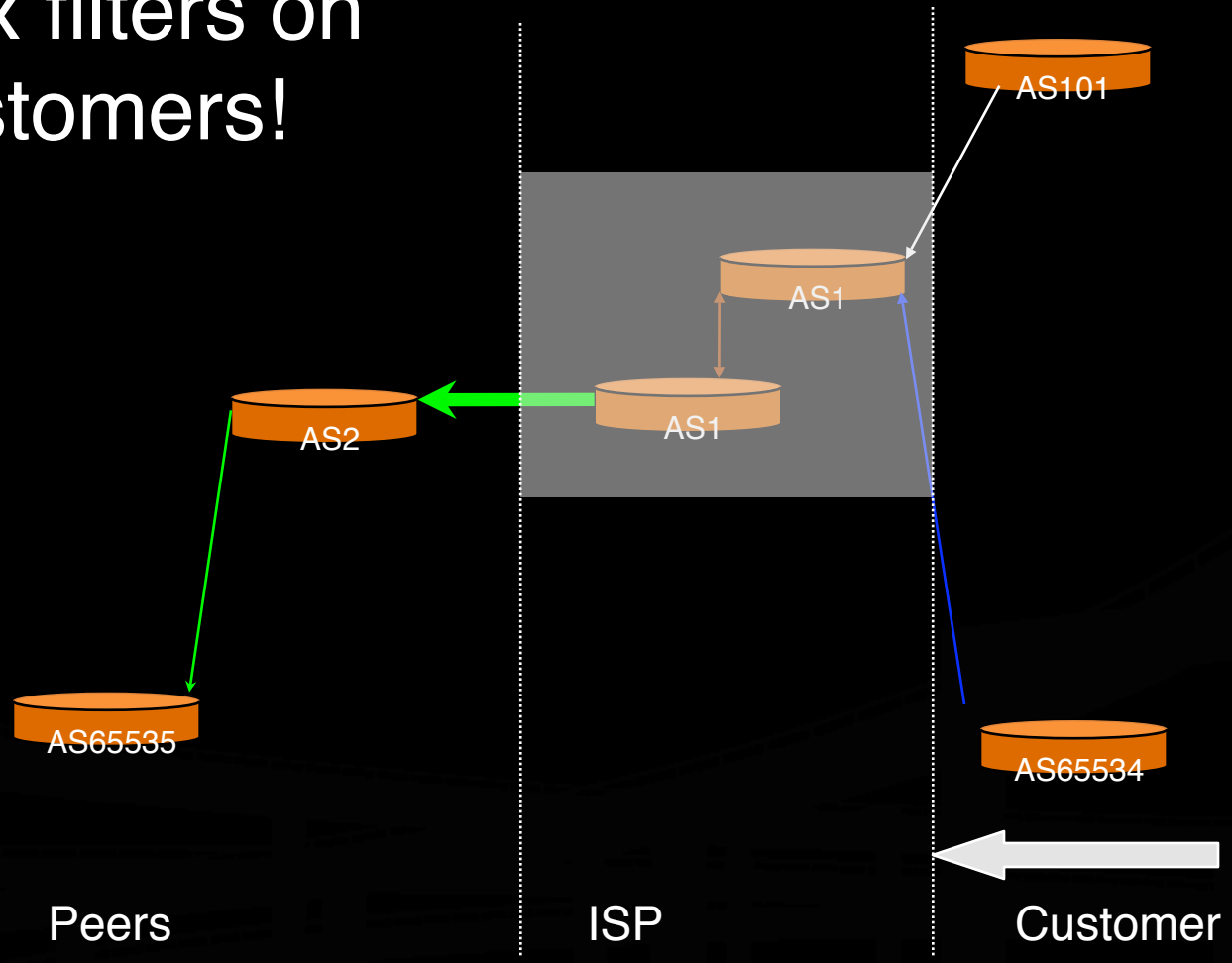
- The bogon list is maintained by the Team Cymru (www.cymru.com)
- While you are at the Team Cymru website, you can also take a look at the secure Router Configuration Template
- You'll find loads of useful information for ISPs and Network people there

Packet Clearing House

Prefix Filters on Customers

Prefix Filters on Customers

- Apply prefix filters on all your customers!



Customer Prefixes

- Service Providers should only accept assigned or allocated prefixes from their downstream peer/customer.
- E.g
  - If the RIR has assigned 202.52.224.0/19 to your customer, accept only that from it
  - If your customer is mulihomed, then accept specific prefix assigned to them by the other ISP, if required

Receiving Customer Prefixes

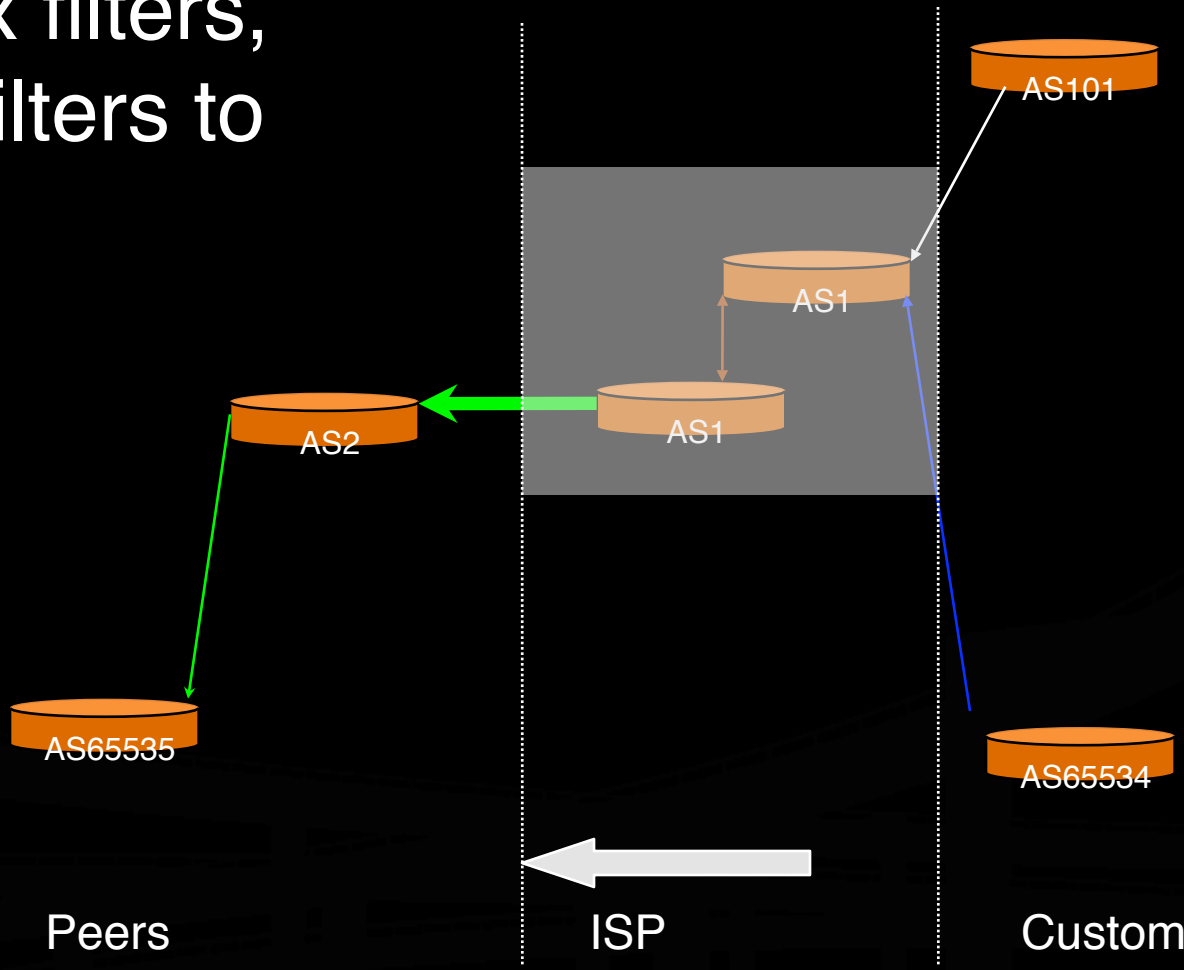
- Configuration example on upstream:

```
router bgp 65534
  neighbor 198.32.231.201 remote-as 201
  neighbor 198.32.231.201 prefix-list CUSTOMER-AS201 in
  !
ip prefix-list CUSTOMER-AS201 permit 201.1.0.0/20
ip prefix-list CUSTOMER-AS201 deny 0.0.0.0/0 le 32
```

Prefixes to Peers

Prefixes to Peers

- Apply prefix filters, AS-PATH filters to your peers!



Packet Clearing House

Prefixes to Peers

- What do you announce to other networks?
  - Your prefixes.
  - Customer's Provider Independent (PI) prefixes
  - More specific customers prefixes (customers who are multihoming)

- What do you *not* send to other network?
  - RFC3330 Prefixes – assume junk will leak into your iBGP.
  - Bogons – assume garbage will leak into your iBGP.
  - Prefixes longer than /24, i.e, /25 to /32

Packet Clearing House

Egress Filter to Peers

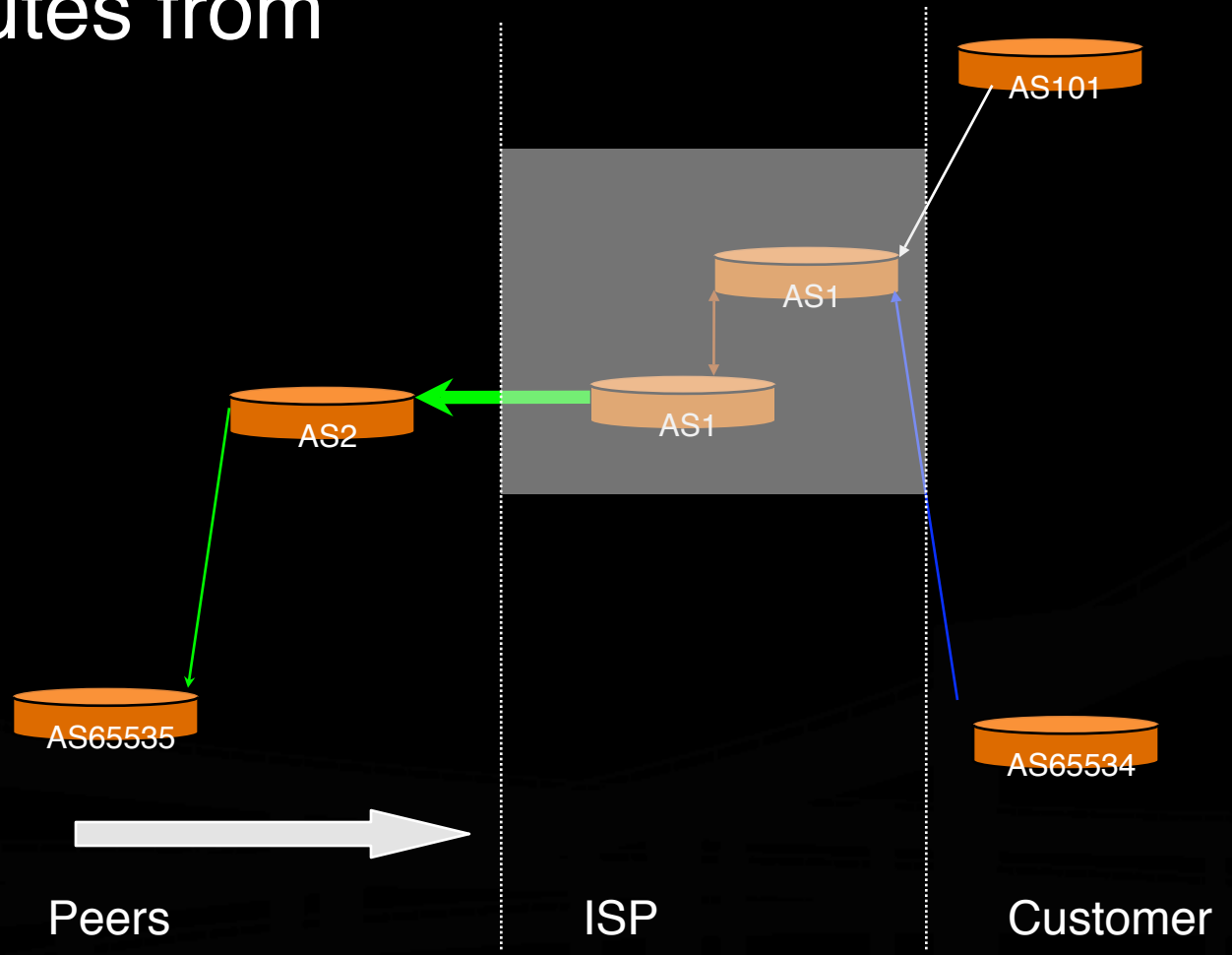
- The egress prefix list can grow to be very large:
  - More specifics for customers.
  - Specific blocks from other ISPs
- AS-PATH filters are more scalable
  - Peer Mailing lists generally send updates about new downstream prefixes

Packet Clearing House

Ingress Prefix Filtering from Peers

Prefixes from Peers

- Filter all routes from your peers!



Packet Clearing House

Filtering Ingress Routes

- Peers and Upstream provide access to the Internet routes
- Ingress filters with Peers
  - The peering policy should have requirements so that filters can be built
  - Max Prefix limit are important
- Ingress filters with Upstream
  - Unless you multihome, full routes are not required, accept only default

PCH
Packet Clearing House

Best Practices for ingress filters

- Don’t accept RFC1918 etc prefixes
- Don’t accept your own prefix
- Don’t accept default (unless you need it)
- Don’t accept prefixes longer than /24
- Consider *Net Police* Filtering

Packet Clearing House

Ingress example - Cisco IOS

```
router bgp 65534
 network 202.52.0.0 mask 255.255.224.0
 neighbor 198.32.231.201 remote-as 201
 neighbor 198.32.231.201 prefix-list in-filter in
!
ip prefix-list in-filter deny 0.0.0.0/0          ! Block default
ip prefix-list in-filter deny 0.0.0.0/8 le 32
ip prefix-list in-filter deny 10.0.0.0/8 le 32
ip prefix-list in-filter deny 127.0.0.0/8 le 32
ip prefix-list in-filter deny 169.254.0.0/16 le 32
ip prefix-list in-filter deny 172.16.0.0/12 le 32
ip prefix-list in-filter deny 192.0.2.0/24 le 32
ip prefix-list in-filter deny 192.168.0.0/16 le 32
ip prefix-list in-filter deny 202.52.0.0/19 le 32 ! Block local prefix
ip prefix-list in-filter deny 224.0.0.0/3 le 32
ip prefix-list in-filter deny 0.0.0.0/0 ge 25            ! Block prefixes >/24
ip prefix-list in-filter permit 0.0.0.0/0 le 32
```

PCH
Packet Clearing House

Packet Filtering Principles

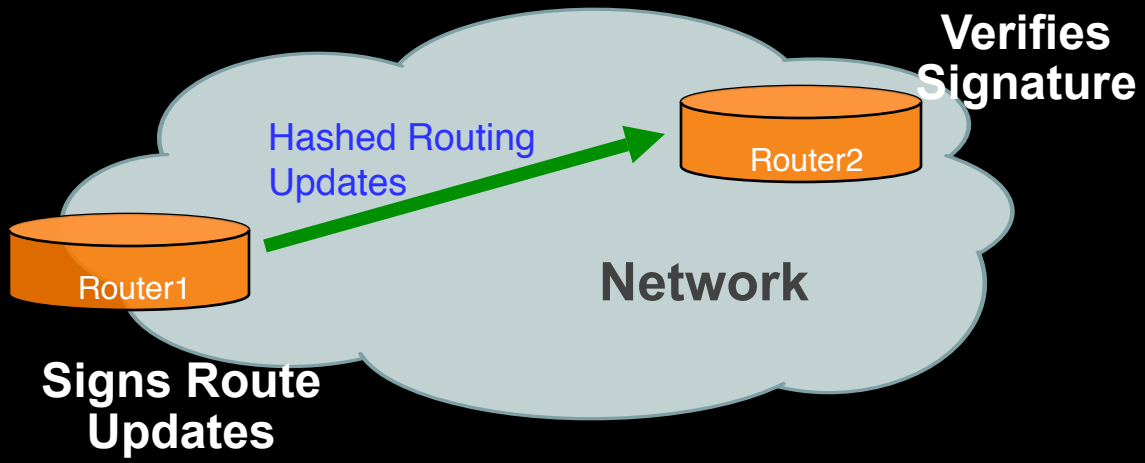
- Filter as close to the edge as possible
- Filter as precisely as possible
- Filter both source and destination where possible

Packet Clearing House

Routing Protocol Updates MD5 Authentication

PCH
Packet Clearing House

Route Authentication



Certifies Authenticity of Neighbor and Integrity of Route Updates

Packet Clearing House

BGP Route Authentication

- Router1 Example

```
router bgp 200
      neighbor 2.2.2.1 remote-as 201
      neighbor 2.2.2.1 description Link to AS-201-Peer
      neighbor 2.2.2.1 password 7 cisco
```

- Router2 Example

```
router bgp 201
      neighbor 2.2.2.2 remote-as 200
      neighbor 2.2.2.2 description Link to AS-200-Peer
      neighbor 2.2.2.2 password 7 cisco
```

Packet Clearing House

Additional BGP Knobs

Packet Clearing House

BGP Maximum Prefix Tracking

- Allow configuration of the maximum number of prefixes a BGP router will receive from a peer
- Two level control
  - Warning threshold: Log warning message
  - Maximum: Tear down the BGP peering, can be setup renew the session after a time period
- Vendor Implementation varies a little bit

Packet Clearing House

BGP Maximum-prefix - Cisco

```
neighbor <x.x.x.x> maximum-prefix <max>
  [<threshold>] [warning-only]
```

- Threshold is an optional parameter between 1 to 100 percent
  - Specify the percentage of <max> that a warning message will be generated; Default is 75%
- Warning-only is an optional keyword which allows log messages to be generated but peering session will not be torn down

PCH
Packet Clearing House

Avoid Default Routes

- ISPs with full BGP feeds should avoid default routes.
- DOS/DDOS attack use spoofed addresses from the un-allocated IPV4 space.
  - See http://www.iana.org/assignments/ipv4-address-space for the latest macro allocations.
- Backscatter traffic from DOS/DDOS targets need to go somewhere. If there is a default, then this traffic will go to this one router and get dropped.
- Dropping backscatter traffic might overload the router.

Packet Clearing House

RFC 2827/BCP 38

Network Ingress Filtering:
Defeating Denial of Service Attacks which employ IP Source Address Spoofing

"Thou shalt only sendth and receiveth IP packets you have rights for"

Packet Clearing House

RFC 2827/BCP 38 Ingress Packet Filtering

- Packets should be sourced from valid, allocated address space, consistent with the topology and space allocation

Packet Clearing House

Guidelines for BCP38

- Networks connecting to the Internet
  - Must use inbound and outbound packet filters to protect network
- Configuration example
  - Outbound—only allow my network source addresses out
  - Inbound—only allow specific ports to specific destinations in

Packet Clearing House

NO BCP38 may mean :

- Devices can (wittingly or unwittingly) send traffic with spoofed and/or randomly changing source addresses out to the network
- Complicates traceback immensely
- Sending bogus traffic is not free

PCH
Packet Clearing House

Techniques for BCP 38 Filtering

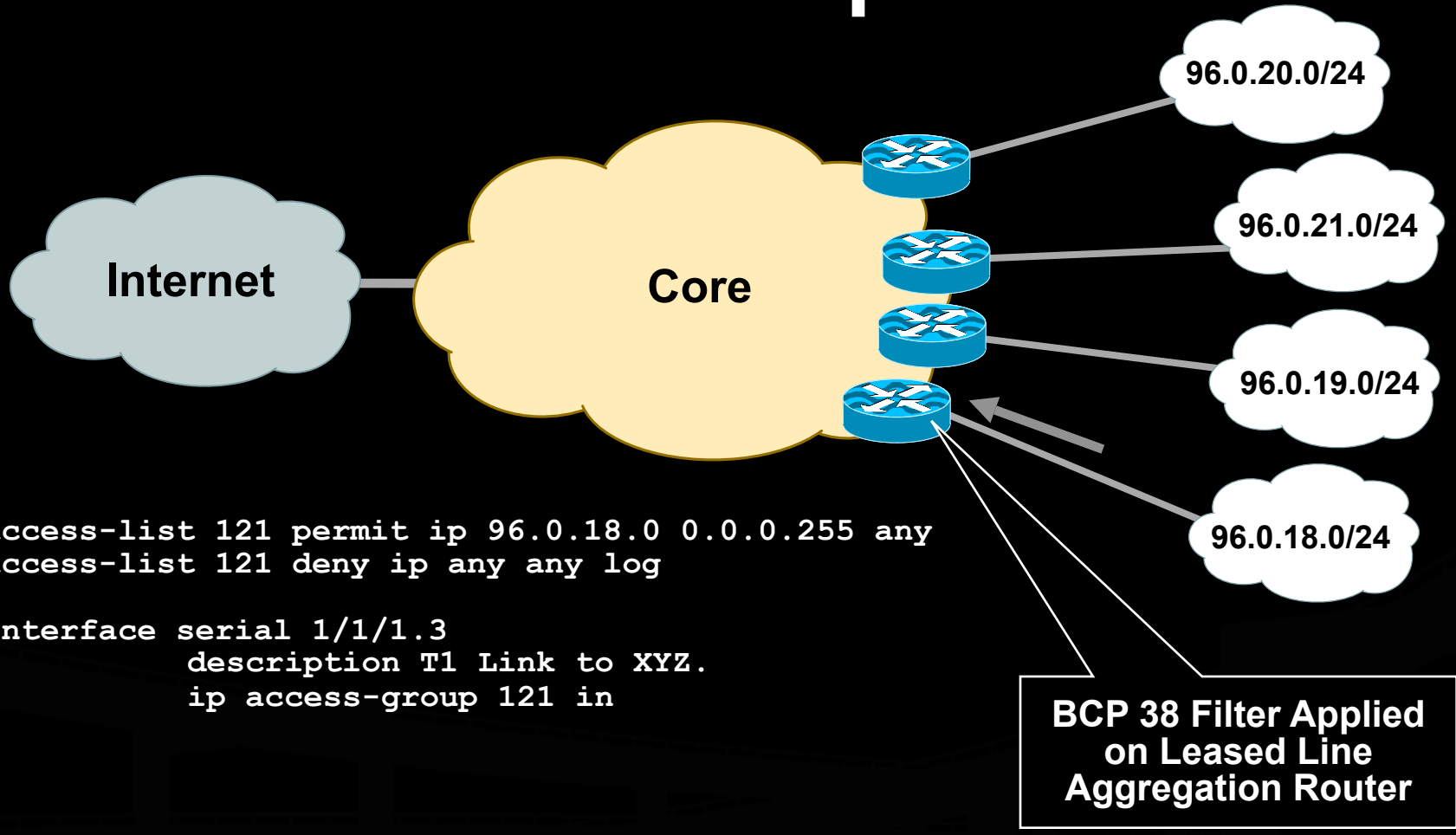
- Static ACLs on the edge of the network
- Dynamic ACLs with AAA profiles
- Unicast RPF strict mode
- IP source guard
- Cable source verify (DHCP)

Packet Clearing House

Using ACLs to Enforce BCP38

- Static ACLs are the traditional method of ensuring that source addresses are not spoofed:
  - Permit all traffic whose source address equals the allocation block
  - Deny any other packet
- Principles:
  - Filter as close to the edge as possible
  - Filter as precisely as possible
  - Filter both source and destination where possible

Static ACL example - Cisco




```
access-list 121 permit ip 96.0.18.0 0.0.0.255 any
access-list 121 deny ip any any log
!
interface serial 1/1/1.3
          description T1 Link to XYZ.
          ip access-group 121 in
!
```

Packet Clearing House

BCP ACL Guidelines

- ISPs
  - Make sure your customers install filters on their routers—give them a template they can use
- Customer end-sites
  - Make sure you install strong filters on routers you use
    to connect to the Internet
  - First line of defense—never assume your ISP will do it

Packet Clearing House

IPv4 Anycast Routing

PCH
Packet Clearing House

What *isn't* Anycast?

- Not a protocol, not a different version of IP, nobody's proprietary technology.
- Doesn't require any special capabilities in the servers, clients, or network.
- Doesn't break or confuse existing infrastructure.

PCH
Packet Clearing House

What *is* Anycast?

- Just a configuration methodology.
- Mentioned, although not described in detail, in numerous RFCs since time immemorial.
- It's been the basis for large-scale content-distribution networks since at least 1995.
- It's gradually taking over the core of the DNS infrastructure, as well as much of the periphery of the world wide web.

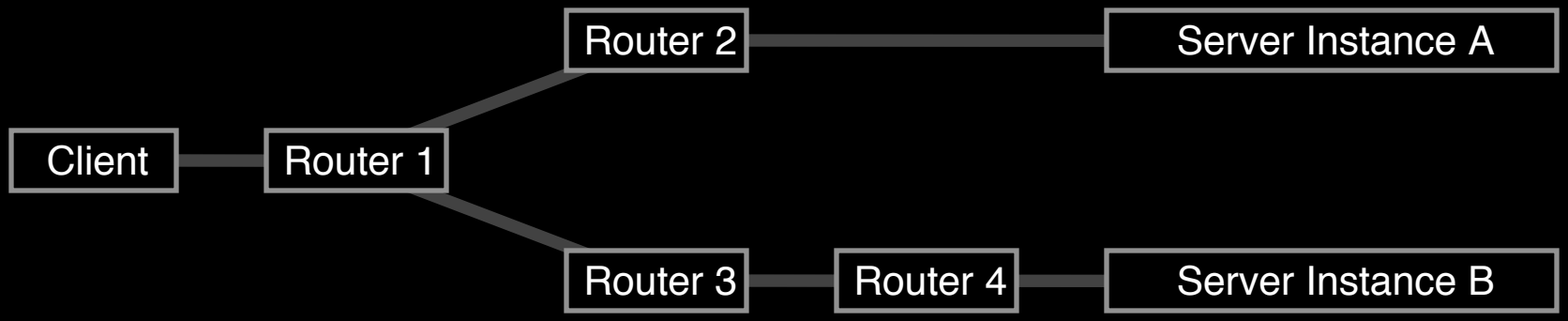
Packet Clearing House

How Does Anycast Work?

- The basic idea is extremely simple:
- Multiple instances of a service share the same IP address.
- The routing infrastructure directs any packet to the topologically nearest instance of the service.
- What little complexity exists is in the optional details.

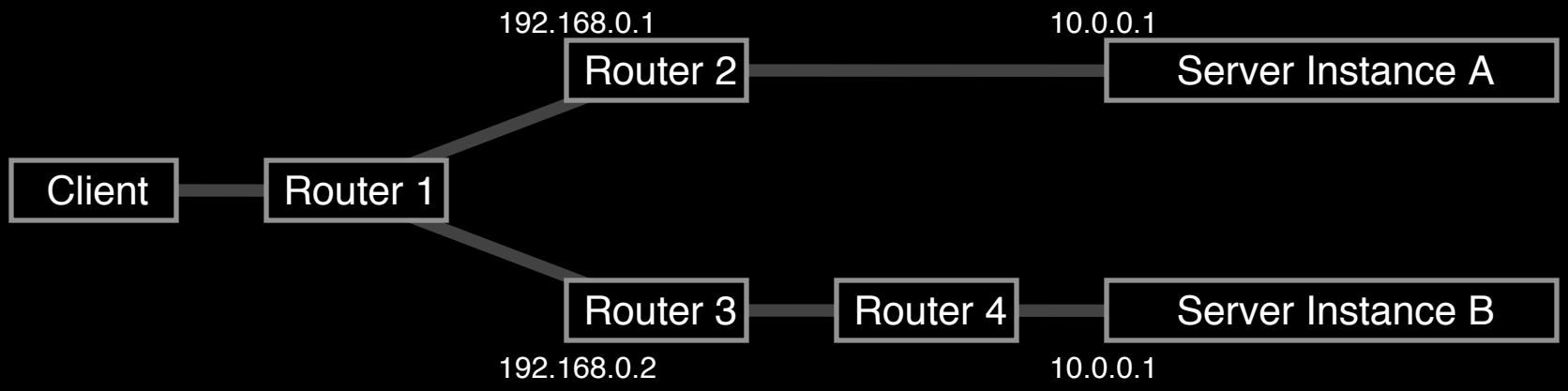
Packet Clearing House

Example



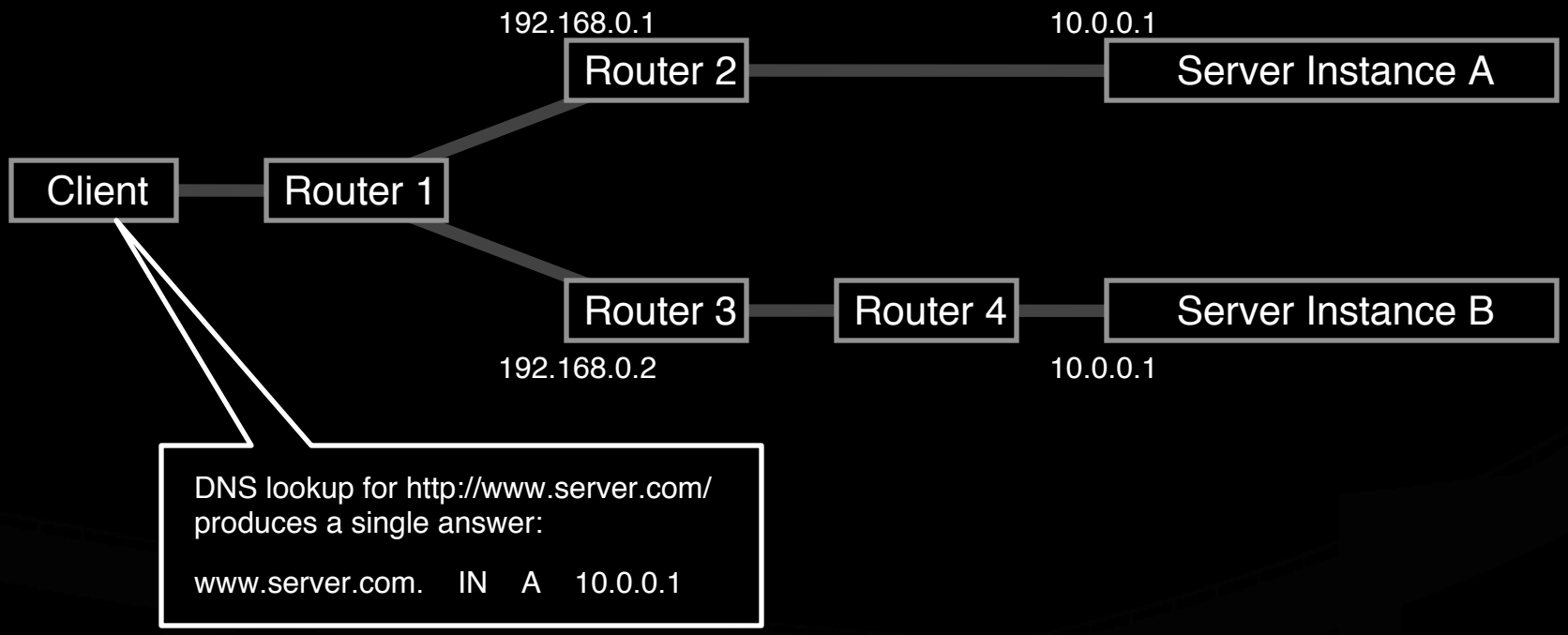
Packet Clearing House

Example



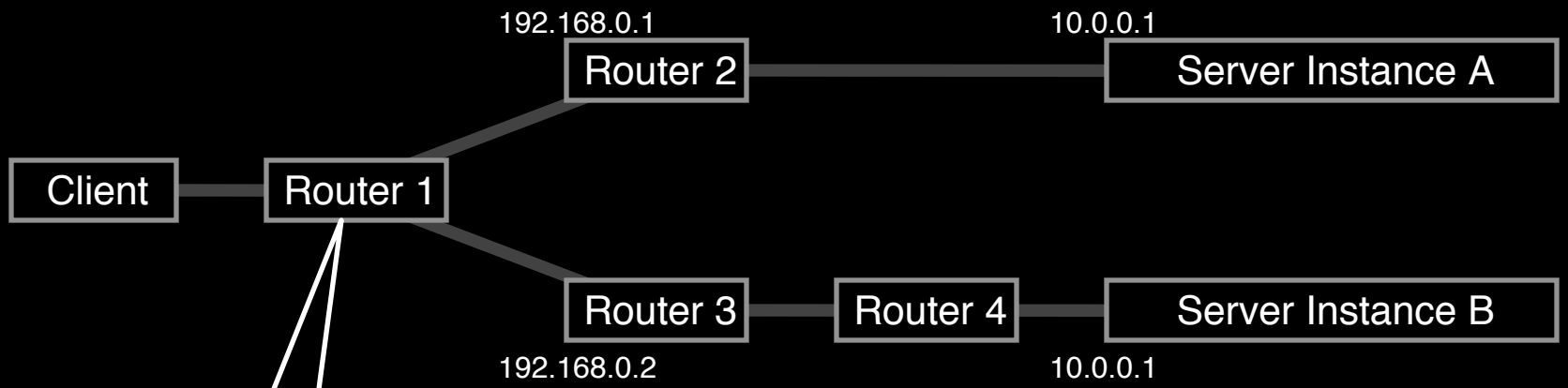
Packet Clearing House

Example



Packet Clearing House

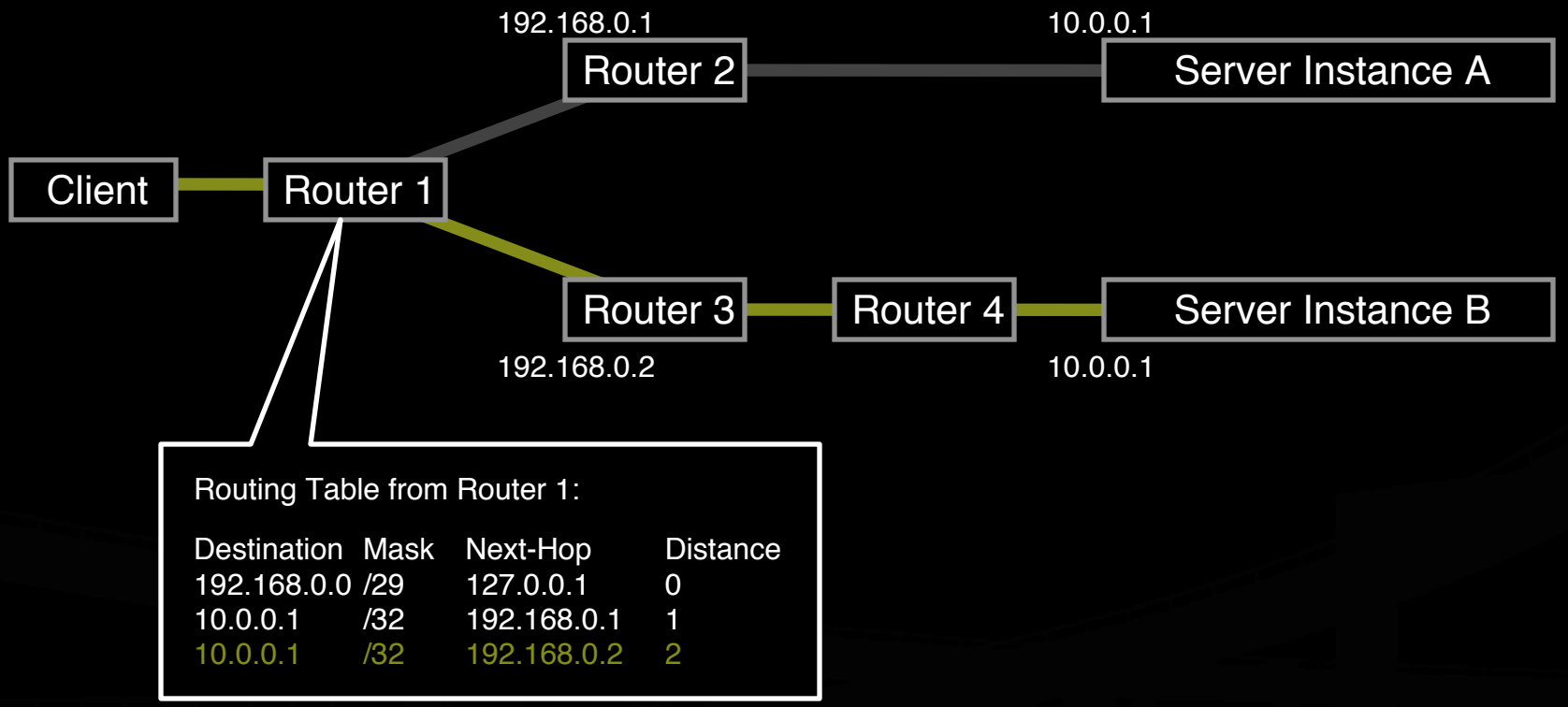
Example



Routing Table from Router 1:

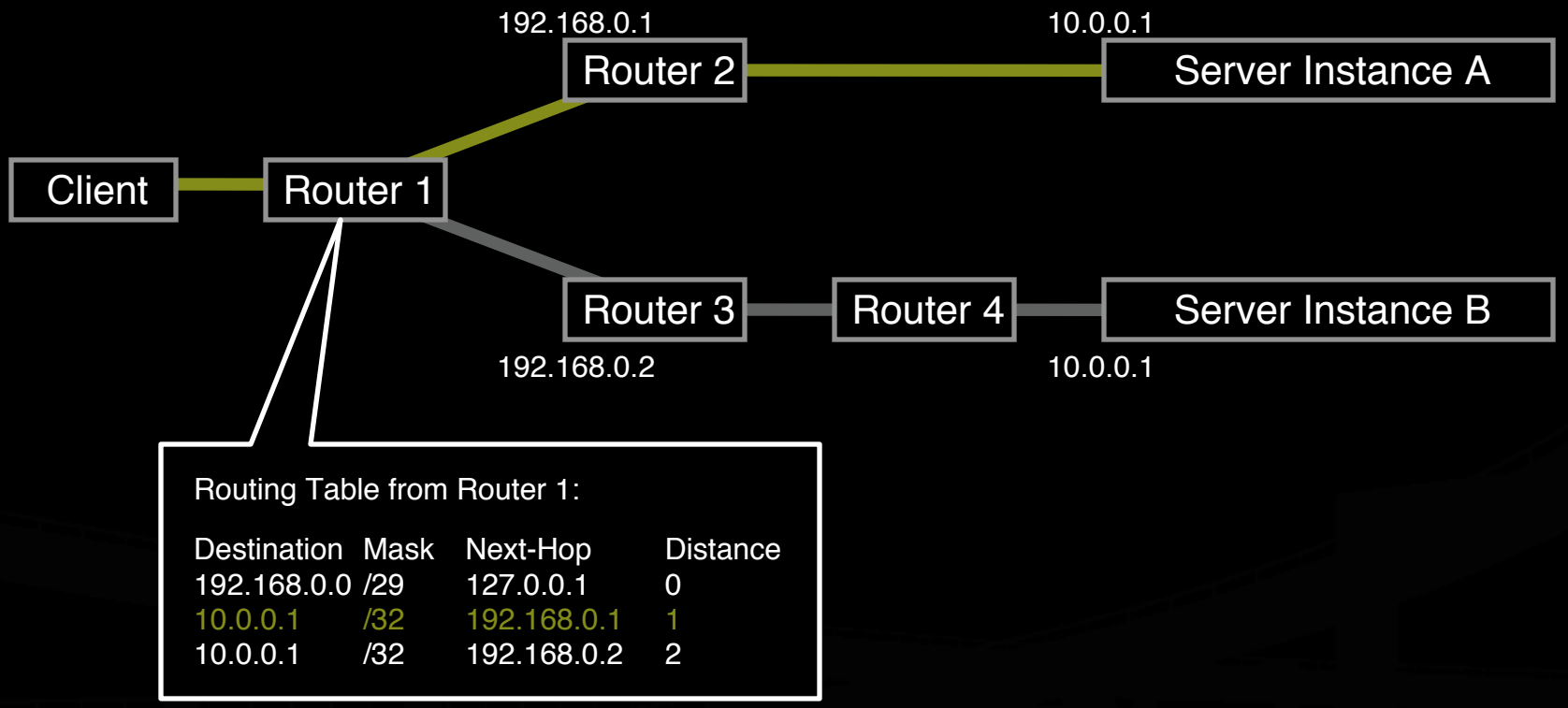
Destination	Mask	Next-Hop	Distance
192.168.0.0	/29	127.0.0.1	0
10.0.0.1	/32	192.168.0.1	1
10.0.0.1	/32	192.168.0.2	2

Example


192.168.0.1
10.0.0.1
Router 2
Server Instance A
Client
Router 1
Router 3
Router 4
Server Instance B
192.168.0.2
10.0.0.1
Routing Table from Router 1:
Destination Mask Next-Hop Distance
192.168.0.0 /29 127.0.0.1 0
10.0.0.1 /32 192.168.0.1 1
10.0.0.1 /32 192.168.0.2 2

Packet Clearing House

Example

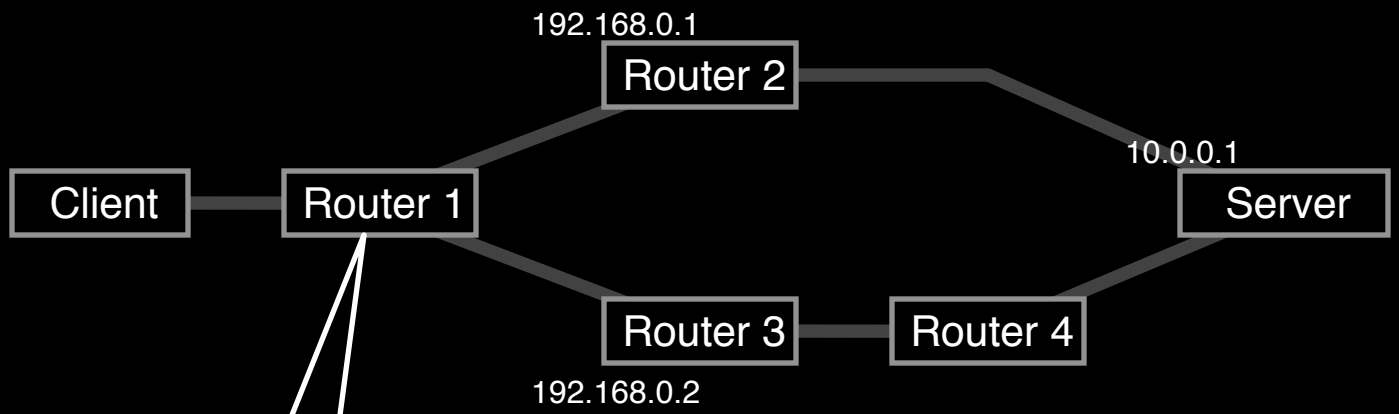


Routing Table from Router 1:

Destination	Mask	Next-Hop	Distance
192.168.0.0	/29	127.0.0.1	0
10.0.0.1	/32	192.168.0.1	1
10.0.0.1	/32	192.168.0.2	2

Example

What the routers think the topology looks like:



Routing Table from Router 1:

Destination	Mask	Next-Hop	Distance
192.168.0.0	/29	127.0.0.1	0
10.0.0.1	/32	192.168.0.1	1
10.0.0.1	/32	192.168.0.2	2

PCH
Packet Clearing House

A Security Ramification

- Anycast server clouds have the useful property of sinking DOS attacks at the instance nearest to the source of the attack, leaving all other instances unaffected.
- This is still of some utility even when DOS sources are widely distributed.

Packet Clearing House

More things to know

Packet Clearing House

Update your whois Information

"Please expect that advanced parts of the community are building filters straight - and strictly - from routing registry data (at least for the RIPE community this is fairly solid data - a benefit which is painfully missing in most other RIR's service areas)."

Ruediger Volk, Deutsche Telecom

(VIX Mailing List, August 17, 05)

PCH
Packet Clearing House

What else should I know ?

- Good ISPs know about their networks, and never block ICMP as a result of some virus activity
- ISPs should not be involved in blocking ports over their network
  - But they can always create a best practice document for their customers not to send internal information on the public internet
- Keep their abuse desk functional and have valid e-mails address for abuse
- Anycast can provide critical service redundancy

PCH
Packet Clearing House

Mailing Lists

- Most countries have regional if not local mailing lists for operational / bogons updates
  - SANOG is South Asia, NANOG is North America, AfNOG is Africa, APOPS is Asia Pacific, EOF is for Europe, Swinog is Swiss, NZNog is Kiwis….. list is long
- Security updates from vendors are sent to most of the list above plus
  - NSP-Sec Mailing List
  - Cisco-NSP, Juniper-NSP and other vendor specific lists
  - CERT mailing list

PCH
Packet Clearing House

Participate

- Networks are only useful when people can use it
- The only way you can make your network work for everyone is by talking to others, so you should participate in these forums.
- If you have an IX, the IX mailing list can be the ideal list for such technical discussion
- Remember, a rouge user anywhere on the Internet can effect you, and chances of attacks originating near your own network is always higher.

Packet Clearing House

INOC DBA

- INOC-DBA: Inter-NOC Dial-by-ASN
- Global Voice-over-IP hotline phone system, directly interconnecting NOCs and SIRTs within carriers, ISPs, exchange points, and vendors.

PCH
Packet Clearing House

How does it work?

- If you just dial an Autonomous System Number, it'll ring a predefined group of phones within that AS. (example: `42`)
- If you dial an ASN and an extension number, it'll ring the phones belonging to that person. (example: `42*WEW`)
- Also, well-known extensions for NOC, abuse, routing, SIRT, et cetera.

Packet Clearing House

Questions ?

Thank You

Gaurab Raj Upadhaya
Peering and Network Group
Packet Clearing House
gaurab@pch.net

With acknowledgements to Philip Fredrick Smith, Rob Thomas, Merike Kaeo and Bill Woodcock

The Best Practices for ISPs tutorial can be found at

http:// www.pch.net / resources / tutorials / ispbcp