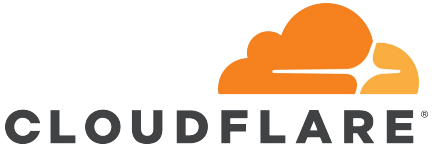


BGP filtering best practice

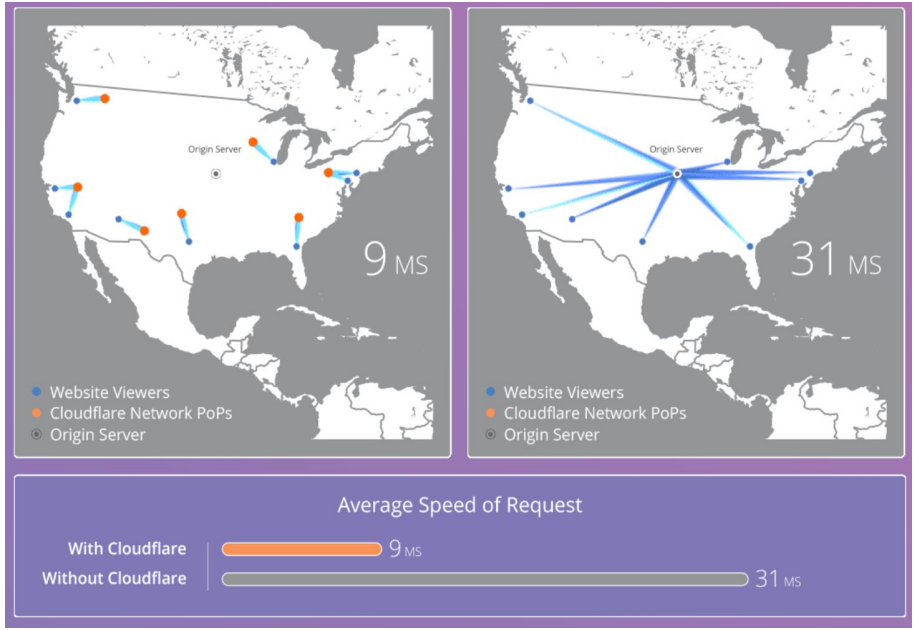
Jimmy Lim
jhalim@cloudflare.com

TWNOG 3
Taipei, 21 June 2019

Cloudflare in a glance

Protect and accelerate any website online

- Direct visitors to nearest entry point
  - Fast
    - Lesser hops
    - Reduced latency
- Save bandwidth
  - Lesser requests to origin
  - Mitigate DDoS
- Resiliency
  - 150+ locations



CLOUDFLARE

CLOUDFLARE

AS Path vs Prefix filtering

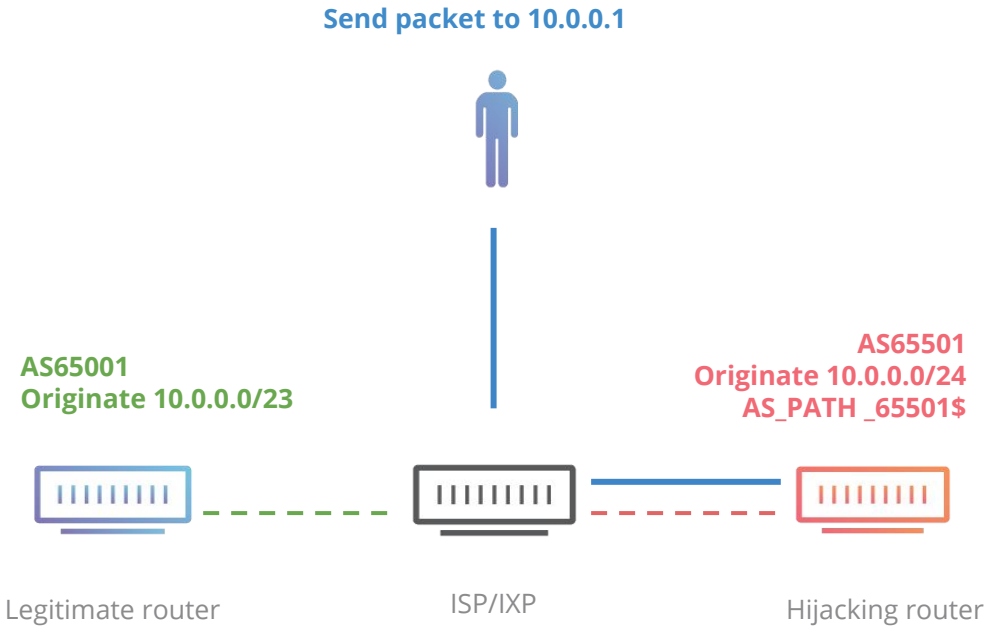
Different type of inbound filtering

- No filtering?
  - No inbound policy filtering
  - Not acceptable
- AS path filtering
  - Filtering based on AS path - a series of autonomous system (AS) numbers with originator's AS number at the end of the path
  - Using regular expression
- Prefix filtering
  - Filtering based on the matching prefixes defined in the prefix list or route filter

Many local IXPs are still doing AS path filtering

- Allow all prefixes that are originated by the ASN
  - No proper validation
- Prone to accidental/wrong announcement
- High risk of prefix hijacking and blackholing

Sample of bad thing about AS Path filtering



CLOUDFLARE

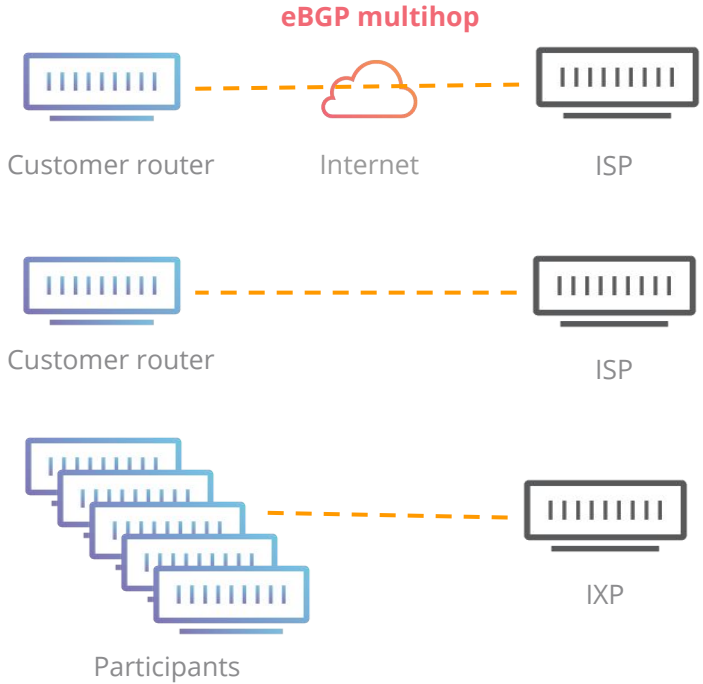
Sample of bad thing about AS Path

- What is so bad about the previous example?

  permit _65501$
- That implies accept any prefixes that are originated by AS65501
  - High chance to create accidental hijacking
  - High risk to hijack or blackhole maliciously

Another sample of bad thing about AS Path filtering

- RTBH feature by ISP/IXPs
  - Allow customers or participants to remotely trigger blackhole to prefixes under their own ASN
    - eBGP multihop blackhole peering with ISPs
    - Direct transit BGP session
    - Using ISP/IXPs blackhole BGP community
  - Accept up to /32



CLOUDFLARE

Why AS Path filtering still exist?

- Difficult to maintain prefix filter
  - Troublesome to validate
- No expertise to configure, too complex?
- IXPs reluctant to do strict filtering
  - Not able to attract all traffic
    - Participants do not update filter regularly/properly

CLOUDFLARE

Prefix list filtering

- Prefix based filtering
  - More validation is done
- Allow up to /24 for IPv4 prefix
- Allow up to /48 for IPv6 prefix
- Automatic update via IRR database

IRR and RPKI

Internet Routing Registry

- Globally distributed routing information database
  - Ensure stability and consistency of Internet-wide routing
  - Sharing information between network operators
- Why use IRR?
  - Route filtering
  - Network troubleshooting
  - Router configuration
  - Global view of Internet routing
- List of IRRs
  - http://www.irr.net/docs/list.html

CLOUDFLARE

AS-SET information in IRR

as-set:	AS-CLOUDFLARE
descr:	Cloudflare, Inc
members:	AS13335
members:	AS3557, AS21556
members:	AS132892, AS133877
members:	AS202623, AS203898
members:	AS394536, AS395747, AS14789
mnt-by:	MNT-CLOUD14
source:	ARIN

CLOUDFLARE

Route object information in IRR

route:	1.1.1.0/24
descr:	Cloudflare, Inc.
origin:	AS13335
mnt-by:	MNT-CLOUD14
notify:	rir@cloudflare.com
remarks:	----------
remarks:	All Cloudflare abuse reporting can be done via
remarks:	https://www.cloudflare.com/abuse
remarks:	----------
source:	ARIN

CLOUDFLARE

Automation to generate prefix filter

```
$ salt edge01.tpe01 irr.gen 8075
edge01.tpe01:
    ----------
    diff:
        [edit policy-options]
        +   policy-statement 4-AS8075-IN {
        +       term FILTER-ROUTES {
        +           from {
        +               route-filter 13.64.0.0/11 upto /24;
        +               route-filter 13.96.0.0/13 upto /24;
        +               route-filter 13.104.0.0/14 upto /24;
        +               route-filter 20.0.0.0/11 upto /24;
        +               route-filter 20.184.0.0/13 upto /24;
        +               route-filter 23.96.0.0/13 upto /24;
        +               route-filter 40.64.0.0/10 upto /24;
        +               route-filter 51.8.0.0/16 upto /24;
        +               route-filter 51.10.0.0/15 upto /24;
        +               route-filter 51.12.0.0/15 upto /24;
        +               route-filter 51.18.0.0/16 upto /24;
        +           }
        +           then accept;
        +       }
        +       term REJECT-ALL {
        +           then {
        +               reject;
        +           }
        +       }
        +   }
```

CLOUDFLARE

IRRs have a very loose security model

- Some database maintainers do not check the authenticity of the entry
  - Records exist within IRRs can be wrong and/or missing
- There are lots of IRRs
  - Mirrors are not always up to date
- No cryptographic signing of records
- Let’s talk about RPKI

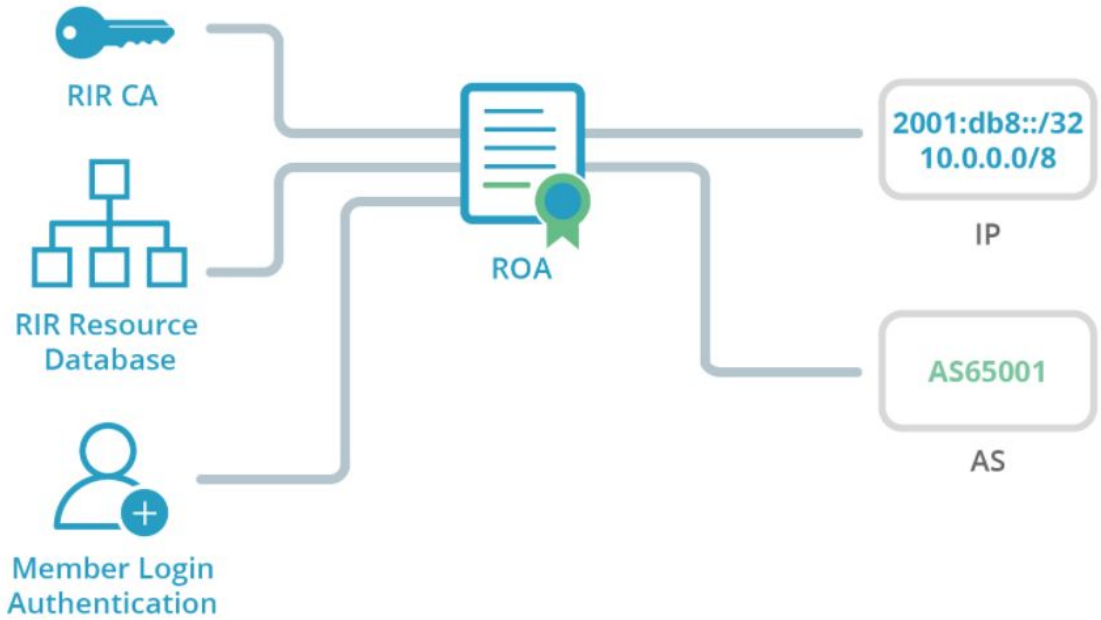
CLOUDFLARE

Resource Public Key Infrastructure

- Cryptographic method of signing records
- Regional Internet Registry (RIR) has a root certificate
  - Generate a signed certificate for Local Internet Registry (network operator)
    - All resources they are assigned with (IPs and ASNs)
- LIR then signs the prefix containing origin AS that they intend to use
  - ROA (Route Object Authorization) is created

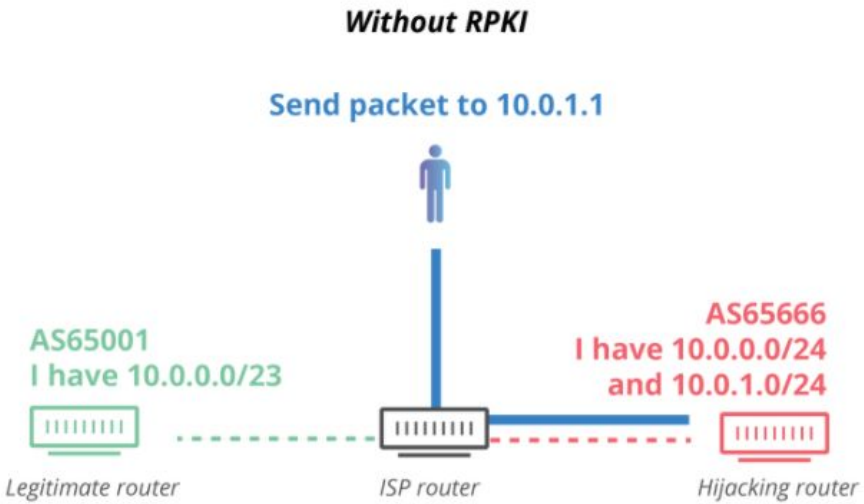
CLOUDFLARE

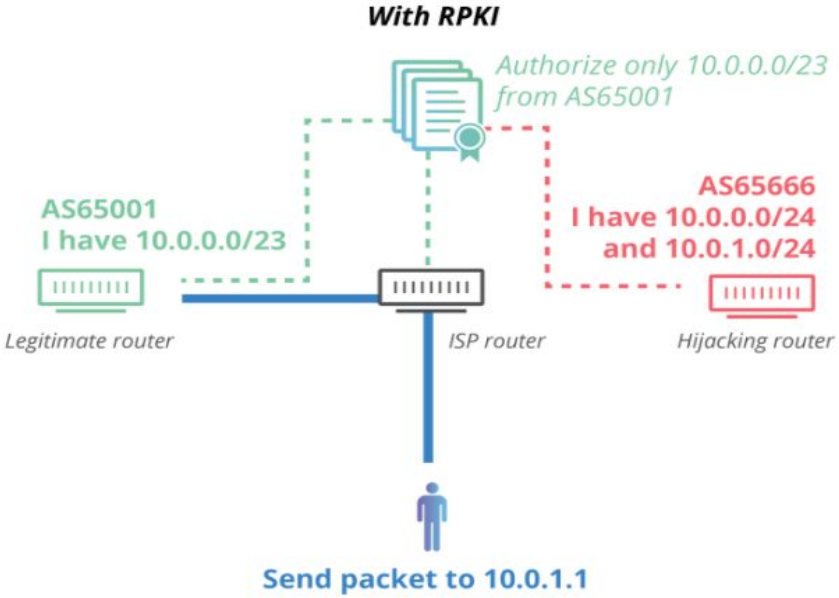
Resource Public Key Infrastructure



CLOUDFLARE

Resource Public Key Infrastructure





CLOUDFLARE

Signing prefixes

- Each LIRs own and manage Internet resources has access to RIR portal
  - Signing their prefixes through the portal or API of their RIR is the easisest way to start with RPKI
- Cloudflare has resources in each of the 5 RIR regions
  - About 800 pefixes announcement over different ASNs
  - We need to ensure the first step is done

CLOUDFLARE

Automation to create ROA

```
$ salt-run rpki.arin_sign $decription 13335 1.1.1.0/24
dict:
    |_
      ----------
      createdDate:
          2019-04-16T22:56:41.296-04:00
    |_
      ----------
      orgHandle:
          CLOUD14
    |_
      ----------
      ticketNo:
          20190416-X525851
    |_
      ----------
      updatedDate:
          2019-04-16T22:56:41.299-04:00
    |_
      ----------
      webTicketStatus:
          IN_PROGRESS
    |_
      ----------
      webTicketType:
          CREATE_ROA
status:
    200
```

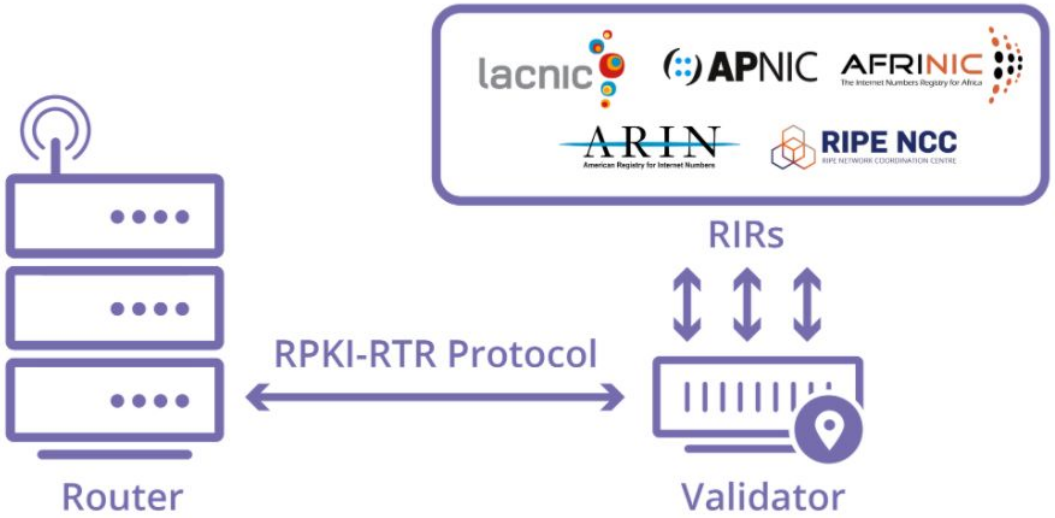
CLOUDFLARE

Enforcing validated prefixes

- Signing the prefixes is one thing
- Ensuring the prefixes we receive match their certificates is another
- Validation is done by synchronizing the RIR databases of ROAs
  - Check the signature of every ROA against the RIR's certification public key
  - Once valid records are known, it is sent to the routers
- Major vendors support a protocol called RPKI to Router Protocol (RTR)
  - A simple protocol for passing a list of valid prefixes with their origin ASN and expected mask length

RPKI to Router Protocol is insecure

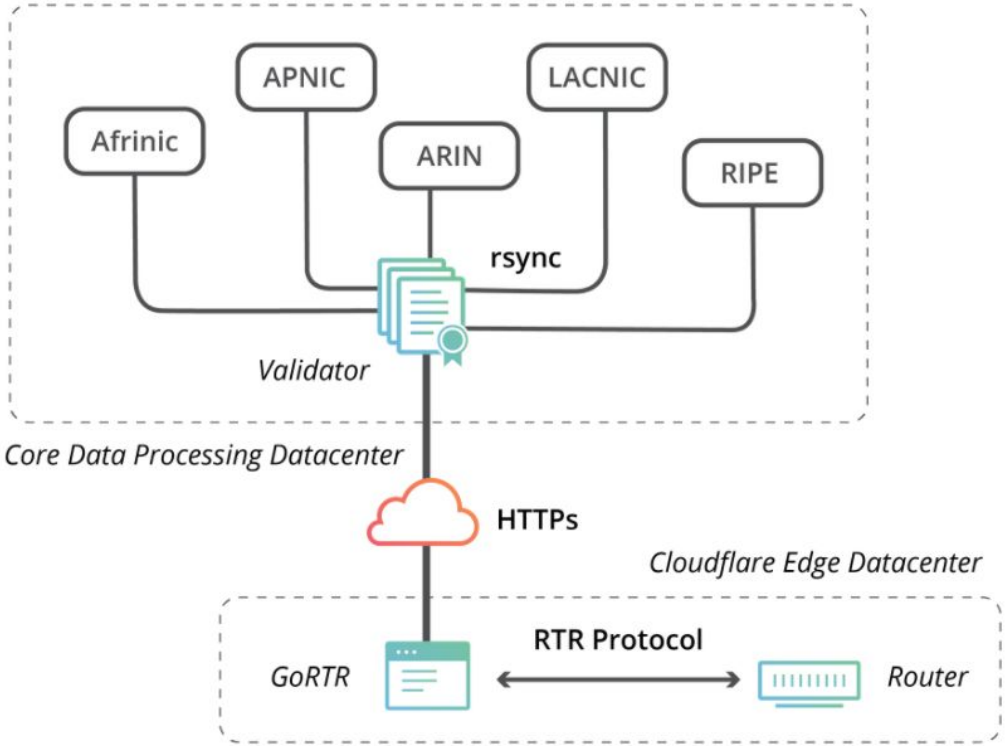
- Vendors implement the insecure transport methods
- Routes sent in clear text over TCP can be tampered with



CLOUDFLARE

Introducing GoRTR

- A lightweight local RTR server
- Distribute it via our own Content Delivery Network
- Fetch the cache file over HTTPS and pass the routes over RTR



CLOUDFLARE

Cloudflare enforcing validated prefixes

```
$ salt edge01.tpe01 state.apply rpki
         diff:
             [edit routing-options]
             +   validation {
             +       group rpki {
             +           session 10.10.10.1 {
             +               port 8282;
             +           }
             +       }
             +   }
```

```
term RPKI-VALID {
    from {
        protocol bgp;
        validation-database valid;
    }
    then {
        validation-state valid;
        next term;
    }
}
term RPKI-INVALID {
    from {
        protocol bgp;›
        validation-database invalid;
    }
    then {
        validation-state invalid;
        reject;
    }
}
```

CLOUDFLARE

Cloudflare enforcing validated prefixes

```
> show validation database origin-autonomous-system 13335
Prefix                 Origin-AS Session                                State
1.0.0.0/24-24              13335 10.10.10.1                           valid
1.1.1.0/24-24              13335 10.10.10.1                           valid
103.22.200.0/23-23         13335 10.10.10.1                           valid
103.22.203.0/24-24         13335 10.10.10.1                           valid
104.16.0.0/20-20           13335 10.10.10.1                           valid
104.16.16.0/20-20          13335 10.10.10.1                           valid
104.16.32.0/20-20          13335 10.10.10.1                           valid
104.16.48.0/20-20          13335 10.10.10.1                           valid
104.16.64.0/20-20          13335 10.10.10.1                           valid
104.16.80.0/20-20          13335 10.10.10.1                           valid
104.16.96.0/20-20          13335 10.10.10.1                           valid
104.16.112.0/20-20         13335 10.10.10.1                           valid
104.16.128.0/20-20         13335 10.10.10.1                           valid
104.16.144.0/20-20         13335 10.10.10.1                           valid
104.16.160.0/20-20         13335 10.10.10.1                           valid
```

CLOUDFLARE

Summary

Wrapping up

- No filtering and AS Path filtering is not acceptable
- Prefix filtering via IRR automation is required
  - Challenge to have this applied in all IXPs
- RPKI is not a replacement yet for IRR
  - Not many has signed their prefixes
  - Not many has enforced validating the prefixes
- Implementing RPKI is achivable
  - Plenty of efforts are needed
  - It is not a bullet proof solution on securing the routing in Internet, but the impact of BGP attacks will be greatly reduced

CLOUDFLARE

Q&A

CLOUDFLARE