Implementation of DNS Anycast

a case study

A. S. M. Shamim Reza | Md. Tawhidul Islam Bhuiyan

Network Operation Center
Link3 Technologies Ltd

bd NOG 9

SANOG XXXII

[~]$whoami

- Linux Geek
- Open Source Software Enthusiast
- EC-Council Certified Security Analyst

ASMShamimReza

shamimreza@link3.net
sohag.shamim@gmail.com

What is Anycast ?

Anycast is a routing method in which incoming requests can be routed to a variety of different locations.

The Journey

- Why Anycast DNS – background history
- Challenges – that we have faced
- Deployment – what we have done
- Configuration – the fun part
- Performance & Security tuning – do’s & don’t

Why Anycast DNS

background history

What we have faced ?

1. Existing DNS server OS version was about to obsolete
2. Resource utilization was always 95%-99%
3. When server was attacked with DDOS
    a. Query response delayed & most of the cases it stopped answering
    b. Unstable DNS service for user internet accesss
4. Log search was not administration friendly
5. No log options for Recursive query

DNS Server – What we had

Software resources	Hardware resources
CentOS 5 32 bit	Core – 2 RAM – 4 GB HDD – Sata 7.2k RPM
bind-utils-9.3.4-10.P1.el5 ypbind-1.19-11.el5 bind-libs-9.3.4-10.P1.el5	

Why we choose Anycast

- Because of the advantages –
  - users of an anycast service will always connect to the closest DNS server; This reduces latency,
  - if one server is being overly loaded, simply deploy another one in a location that would allow it to take some proportion of the overloaded server's requests; horizontal scaling.

- We need to have 1 single IP for the Recursive DNS server all over Bangladesh.
- As we are also expanding our network infrastructure, we didn’t want our zonal internet user to be depended on our Central Data Center based DNS system.

Challenges

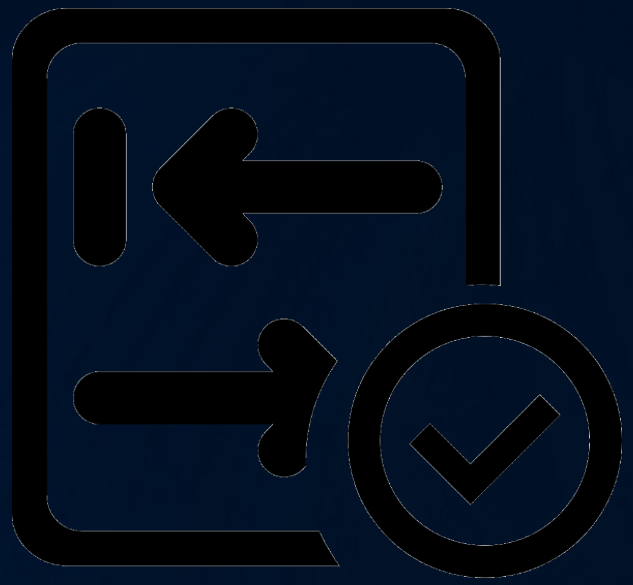
that we have faced

Technical Difficulties - Not really but

- Monitoring was more complicated
- Monitoring the anycast IP can not be done centrally
- Changing the DNS server IP of all the internet users
  - Informed client with email, sms and other notification option

Deployment

what we have done



Decision – we have taken

- Security first
- Deploy with updated OS
- Divide the Authoritative & Recursive in to TWO server
- Deploy the IP Anycast for Recursive DNS only
- Configure the caching log based on search criteria
- Agent based Central Monitoring to monitor individual servers

Procedure – we have listed

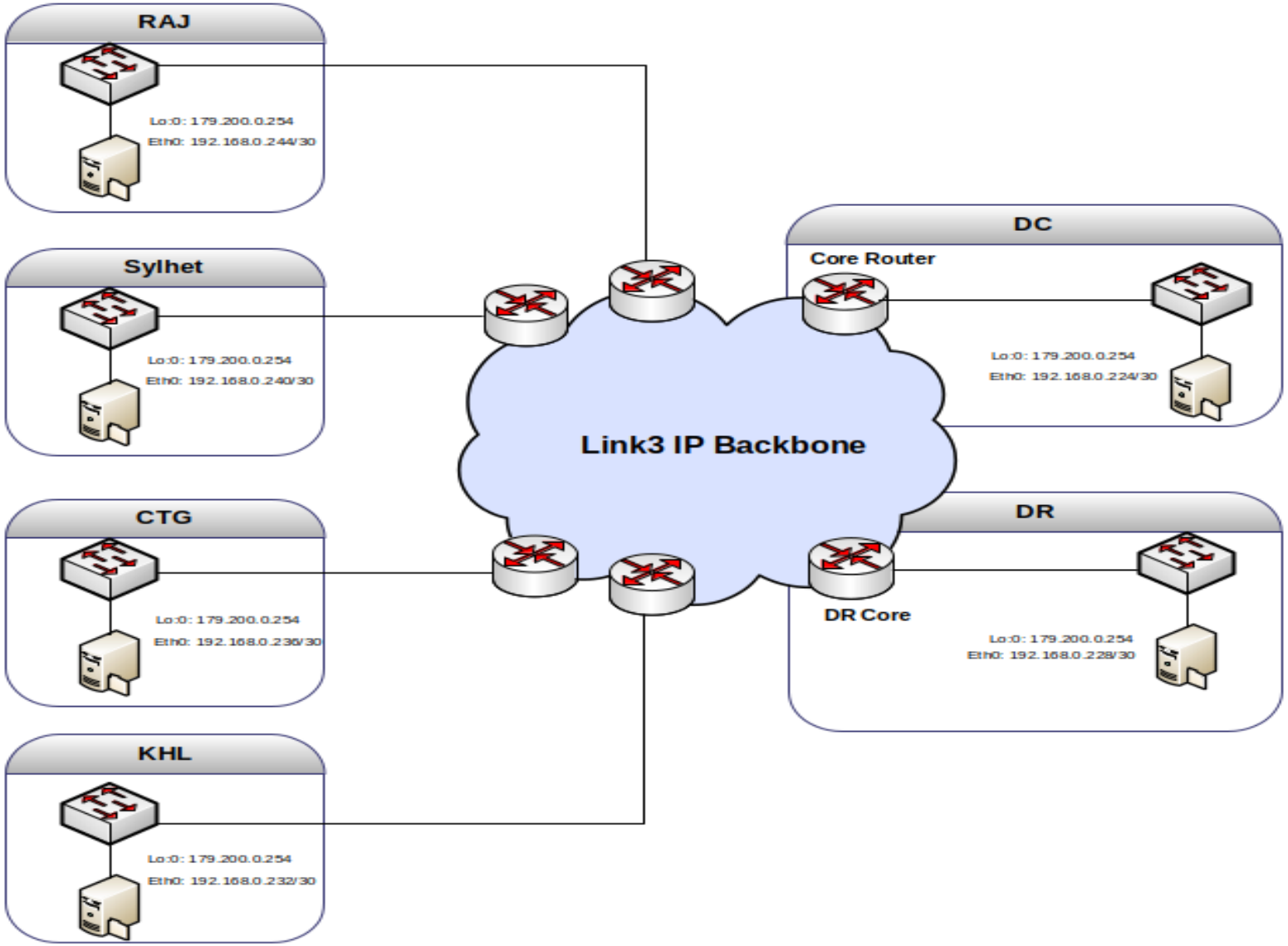
- Address selection
- Host configuration
- Service configuration
- Network configuration
- Follow standard security measures

Resources – that we have allocated for server

Software Resources	Hardware Resources
CentOS 7.5 64 bit	CPU Core – 4 with 2 Socket RAM – 4 GB DDR4 HDD – Sata SAS 15k RPM
rpcbind-0.2.0-44.el7.x86_64 bind-chroot-9.9.4-61.el7.x86_64 bind-license-9.9.4-61.el7.noarch bind-utils-9.9.4-61.el7.x86_64 bind-9.9.4-61.el7.x86_64 bind-libs-lite-9.9.4-61.el7.x86_64 bind-libs-9.9.4-61.el7.x86_64	
iptables-1.4.7-16.el6.x86_64 iptables-ipv6-1.4.7-16.el6.x86_64	
quagga-0.99.22.4-5.el7_4.x86_64	

Network Diagram

ANYCAST DNS INFRASTRUCTURE



With the New System – challenges

- Query response was slower/ Some of the users are not getting response
- Server resources was filled up with log files and DNS service was stacked, but BGP was up; so no one was getting internet and the anycast shifting didn’t happened.

With the New System – why we suffered

- Performance tuning wasn’t done
- Monitoring wasn’t placed properly
- Query hit increased to 6k/second

With the New System – recovery steps

- Configured the log rotation based on file size
- Decided to move all the log to the central server after every one hour
- Write up a script to sense dns service;
  - if PiD is null value then shutdown the BGP. That will automatically shift the IP Anycast to nearest one.
  - If PiD is ok then check with localhost if it answers to DNS query, if not then shutdown the BGP.

With the New System – the script

```
#!/bin/bash
DNSUP=`/usr/bin/dig @179.100.0.254 localhost. A +short`
if [ "$DNSUP" != "127.0.0.1" ];
then
echo "Stopping Anycast...."
   /etc/init.d/bgpd stop
   /etc/init.d/zebra stop
        echo "Stopped: DC Anycast DNS has stopped working, BGP has already been shutdown, Please check the system right now." | mailx -S smtp=smtp.notification.net:25 -s "Alert: Stopped - DC Anycast DNS has stooped working" nothing@notification.com
else
    echo "Everything's good... Do nothing..."
fi
```

Configuration

the fun part

Configuration - address selection

- Dedicated unique management IP for each host
- Designated 1 single /32 for Anycast address for all servers
- Private ASN 65430 for peering with ISP core

Dhaka Server - assigned anycast address

Anycast address as an additional loopbacks

[root@dc-anycast-dns network-scripts]# ifconfig lo:0
 	
lo:0: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
  inet 179.100.0.254 netmask 255.255.255.255
  loop txqueuelen 1 (Local Loopback)

Dhaka Server - named service

Configuring named service to listen on anycast address

```
[root@dc-anycast-dns etc]# vim /var/named/chroot/etc/named.conf
options {
        listen-on port 53 { 127.0.0.1; 179.100.0.254; };
        directory       "/var/named";
        dump-file       "/var/named/data/cache_dump.db";
        statistics-file "/var/named/data/named_stats.txt";
        memstatistics-file "/var/named/data/named_mem_stats.txt";
        allow-query     { localhost; 192.168.0.0/16; };
        allow-query-cache       { localhost; 192.168.0.0/16; };
        allow-recursion { localhost; 192.168.0.0/16;  };
        version "go to sleep" ;
        recursive-clients 100000;
};
```

Dhaka Server - named service

Configuring named service for separate query logging

```
logging {
   channel default_file {
      file "/var/named/chroot/var/log/named/default.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
channel queries_file {
      file "/var/named/chroot/var/log/named/queries.log" versions 2 size 4096m;
      severity dynamic;
      print-time yes;
      };
channel resolver_file {
      file "/var/named/chroot/var/log/named/resolver.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
channel security_file {
      file "/var/named/chroot/var/log/named/security.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
   category default { default_file; };
   category security { security_file; };
   category resolver { resolver_file; };
   category queries { queries_file; };
     };
```

Dhaka Server - quagga & bgp

Configuring zebra.conf

```
[root@dc-anycast-dns quagga]# # vim /etc/quagga/zebra.conf

hostname dc-anycast-dns.link3.net
!
enable password NothingToSay
!
interface eth0
ip address 192.168.0.226/30
!
interface lo:0
ip address 179.200.0.254/32
!
interface lo
!
line vty
!
```

Dhaka Server - quagga & bgp

Configuring bgpd.conf

```
[root@dc-anycast-dns quagga]# vim /etc/quagga/bgpd.conf
hostname dc-anycast-dns.link3.net
password NothingToSay
log stdout
!
router bgp 65430
 network 179.200.0.254/32
 neighbor 192.168.0.225 remote-as 23688
 neighbor 192.168.0.225 description BTS
 neighbor 192.168.0.225 activate
 neighbor 192.168.0.225 next-hop-self
 neighbor 192.168.0.225 remove-private-AS
 neighbor 192.168.0.225 soft-reconfiguration inbound
 neighbor 192.168.0.225 prefix-list anycast out
 neighbor 192.168.0.225 prefix-list default in
!
ip prefix-list default seq 15 permit 0.0.0.0/0
ip prefix-list anycast seq 5 permit  179.200.0.254/32
```

Dhaka Router - announcing route

Configuring BGP from router
router bgp 23688 network 192.168.0.224 mask 255.255.255.252 neighbor 192.168.0.226 remote-as 65430 neighbor 192.168.0.226 description DC-DNS_Anycast-SERVER neighbor 192.168.0.226 activate neighbor 192.168.0.226 next-hop-self neighbor 192.168.0.226 default-originate neighbor 192.168.0.226 remove-private-as neighbor 192.168.0.226 soft-reconfiguration inbound neighbor 192.168.0.226 prefix-list anycast-DNS-in in neighbor 192.168.0.226 prefix-list default out ip prefix-list anycast-DNS-in seq 10 permit 179.200.0.254/32 ip prefix-list default seq 5 permit 0.0.0.0/0

Sylhet Server - assigned anycast address

Anycast address as an additional loopbacks

[root@syl-anycast-dns network-scripts]# ifconfig lo:0
 	
lo:0: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
  inet 179.100.0.254 netmask 255.255.255.255
  loop txqueuelen 1 (Local Loopback)

Sylhet Server - named service

Configuring named service to listen on anycast address

```
[root@syl-anycast-dns etc]# vim /var/named/chroot/etc/named.conf
options {
        listen-on port 53 { 127.0.0.1; 179.100.0.254; };
        directory       "/var/named";
        dump-file       "/var/named/data/cache_dump.db";
        statistics-file "/var/named/data/named_stats.txt";
        memstatistics-file "/var/named/data/named_mem_stats.txt";
        allow-query     { localhost; 192.168.0.0/16; };
        allow-query-cache       { localhost; 192.168.0.0/16; };
        allow-recursion { localhost; 192.168.0.0/16;  };
        version "go to sleep" ;
        recursive-clients 100000;
};
```

Sylhet Server - named service

Configuring named service for separate query logging

```
logging {
   channel default_file {
      file "/var/named/chroot/var/log/named/default.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
channel queries_file {
      file "/var/named/chroot/var/log/named/queries.log" versions 2 size 4096m;
      severity dynamic;
      print-time yes;
      };
channel resolver_file {
      file "/var/named/chroot/var/log/named/resolver.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
channel security_file {
      file "/var/named/chroot/var/log/named/security.log" versions 2 size 200m;
      severity dynamic;
      print-time yes;
      };
   category default { default_file; };
   category security { security_file; };
   category resolver { resolver_file; };
   category queries { queries_file; };
     };
```

Sylhet Server - quagga & bgp

Configuring zebra.conf

```
[root@syl-anycast-dns quagga]# # vim /etc/quagga/zebra.conf

hostname sylt-anycast-dns.link3.net
!
enable password NothingToSay
!
interface eth0
ip address 192.168.0.232/30
!
interface lo:0
ip address 179.200.0.254/32
!
interface lo
!
line vty
!
```

Sylhet Server - quagga & bgp

Configuring bgpd.conf

```
[root@syl-anycast-dns quagga]# vim /etc/quagga/bgpd.conf
hostname sylt-anycast-dns.link3.net
password NothingToSay
log stdout
!
router bgp 65430
 network 179.200.0.254/32
 neighbor 192.168.0.233 remote-as 23688
 neighbor 192.168.0.233 description BTS
 neighbor 192.168.0.233 activate
 neighbor 192.168.0.233 next-hop-self
 neighbor 192.168.0.233 remove-private-AS
 neighbor 192.168.0.233 soft-reconfiguration inbound
 neighbor 192.168.0.233 prefix-list anycast out
 neighbor 192.168.0.233 prefix-list default in
!
ip prefix-list default seq 15 permit 0.0.0.0/0
ip prefix-list anycast seq 5 permit  179.200.0.254/32
```

Sylhet Router - announcing route

Configuring BGP from router

```
router bgp 23688
 network 192.168.0.234 mask 255.255.255.252
 neighbor 192.168.0.234 remote-as 65430
 neighbor 192.168.0.234 description Sylt-DNS_Anycast-SERVER
 neighbor 192.168.0.234 activate
 neighbor 192.168.0.234 next-hop-self
 neighbor 192.168.0.234 default-originate
 neighbor 192.168.0.234 remove-private-as
 neighbor 192.168.0.234 soft-reconfiguration inbound
 neighbor 192.168.0.234 prefix-list anycast-DNS-in in
 neighbor 192.168.0.234 prefix-list default out
ip prefix-list anycast-DNS-in seq 10 permit 179.200.0.254/32
ip prefix-list default seq 5 permit 0.0.0.0/0
```

Performance and security tuning

do’s & don’t

Performance tuning

- Checked the System –
  - # /sbin/sysctl net.netfilter.nf_conntrack_count

    net.netfilter.nf_conntrack_count = 262144
- Changed it –
  - # sysctl -w net.netfilter.nf_conntrack_max=524288

Security Measures – that has been taken

- Install & Configure the named service with least privileges **CHROOT**
  - [root@bd-anycast-dns quagga]# cd /var/named/chroot/ && ls

    dev etc run usr var

- Hide the bind version
  - [root@bd-anycast-dns etc]# cat /var/named/chroot/etc/named.conf

    version “please don’t ask my name" ;

- Restrict queries
  - [root@bd-anycast-dns etc]# cat /var/named/chroot/etc/named.conf

    allow-query { localhost; 192.168.0.0/16; };
    allow-query-cache { localhost; 192.168.0.0/16; };

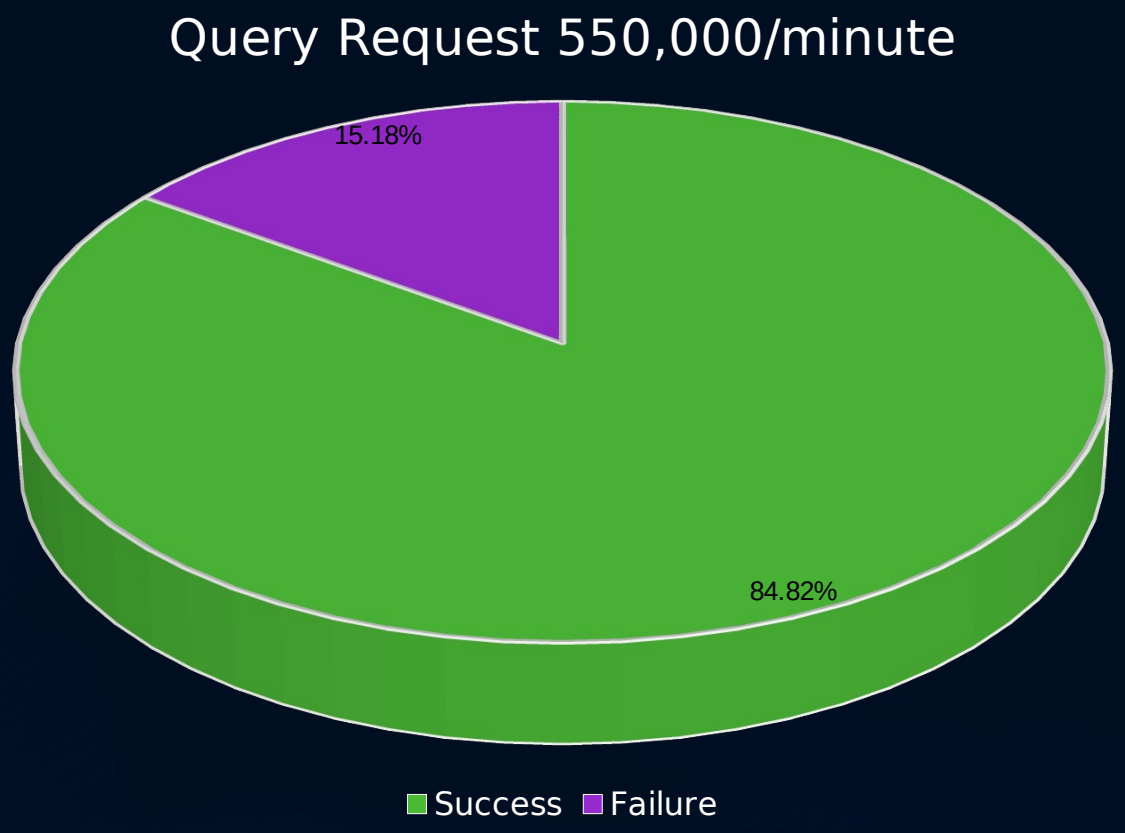
    allow-recursion { localhost; 192.168.0.0/16; };

- Named service was configured to Listen to only Anycast Address
  - [root@bd-anycast-dns etc]# cat /var/named/chroot/etc/named.conf

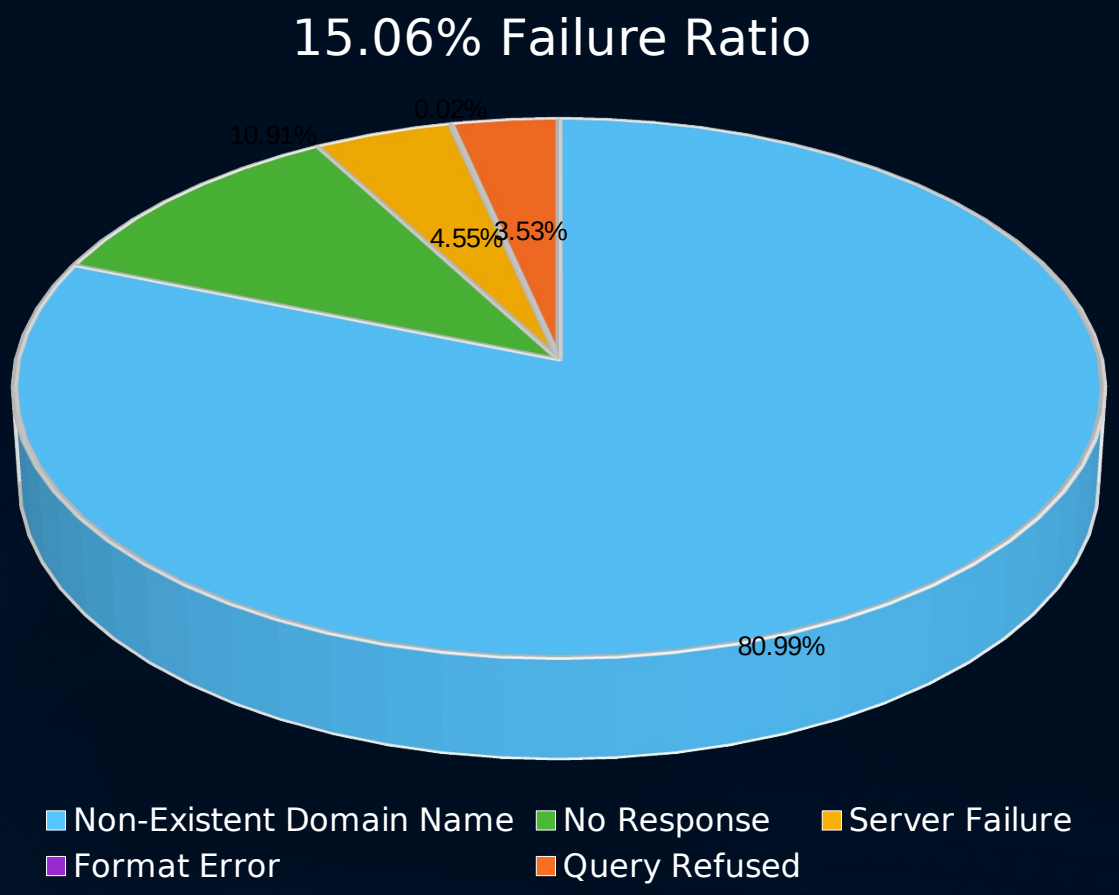
    listen-on port 53 { 127.0.0.1; 179.100.0.254; };

DNS service analysis

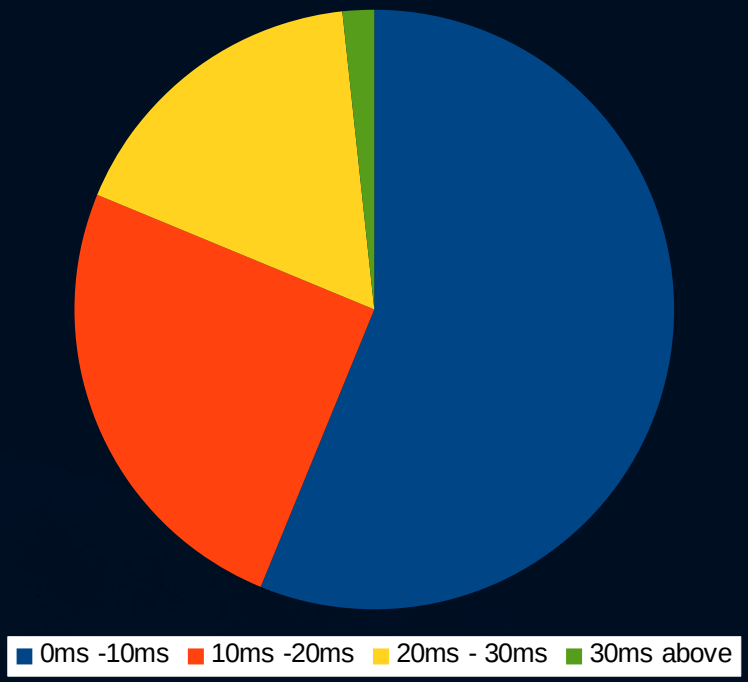
Success and Failure Ratio



DNS Failure Reasons



DNS Resolution Time



Thank You!