IP Anycast for Recursive DNS Service

A. S. M. Shamim Reza

Deputy Manager
Network Operation Center
Link3 Technologies Limited

SANOG 34

[~]$whoami

– 10 years, working for Link3 Technologies Limited
– Opensource Software Enthusiast
– InfoSec Professional
– Passionate about Artificial Intelligence

EC-Council Certified Security Analyst

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

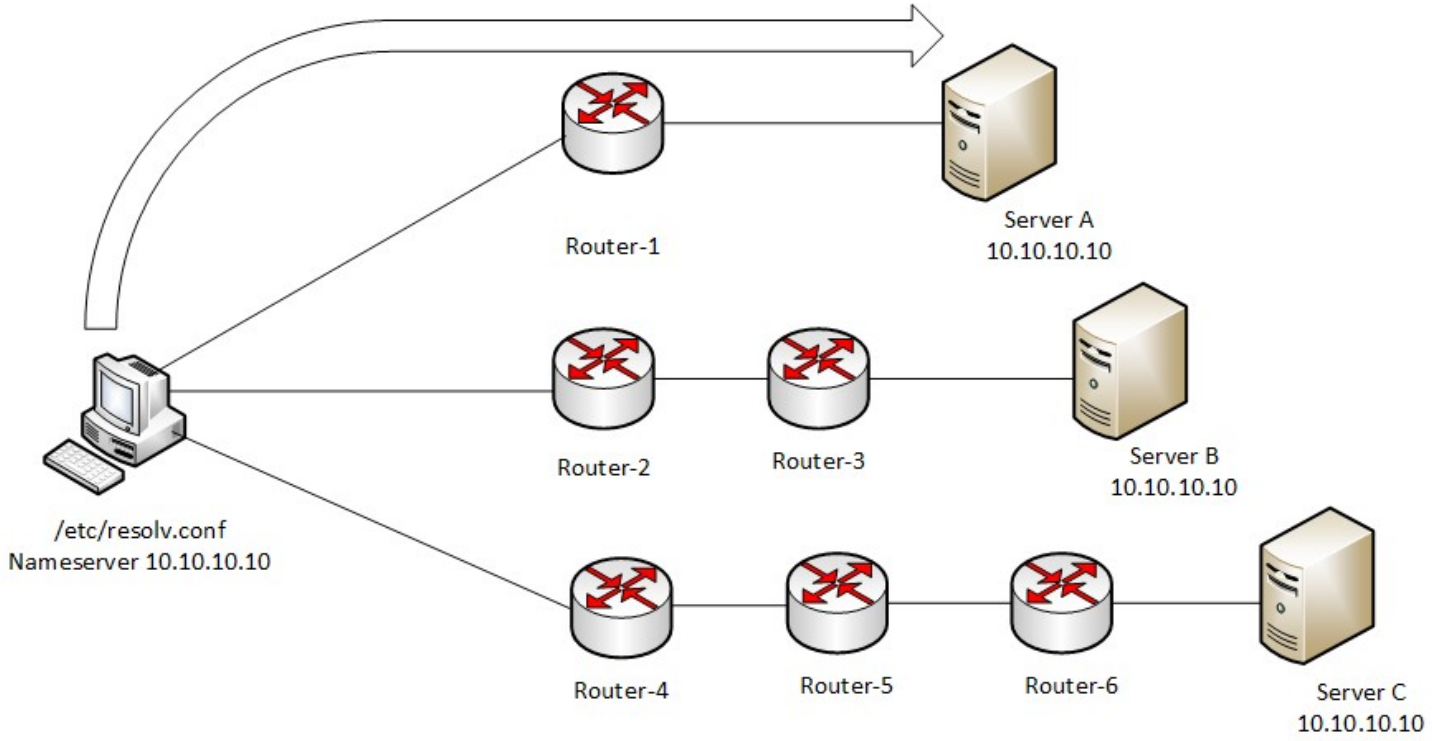
@*asmshamimreza* on **Linkedin**

Agenda

– How to deploy IP Anycast for Recursive DNS service.

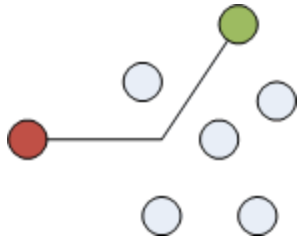
What is Anycast ?

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

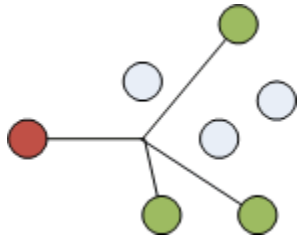


What is Anycast ?

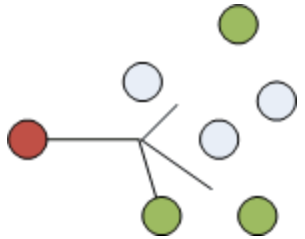
– Unicast refers to 1-to-1 conversations



– Multicast refers to 1-to-many conversations



– Anycast refers to 1-to-any conversations.



What isn’t Anycast ?

– Not a protocol, not a different version of IP.

– Doesn’t require any special capabilities in the servers, clients, or network.

– Doesn’t break or confuse existing infrastructure.

Why Anycast is Important?

- Anycast is designed for short queries.
- Connection-less protocols are ideal.
- Anycast provides a level of redundancy that DNS round-robin cannot provide.
- The overall service can be insulated somewhat from Denial of Service Attacks
- The service scales very well.
- Traffic is routed to the server that is the closest (best path).

How Anycast works ?

The basic idea is –

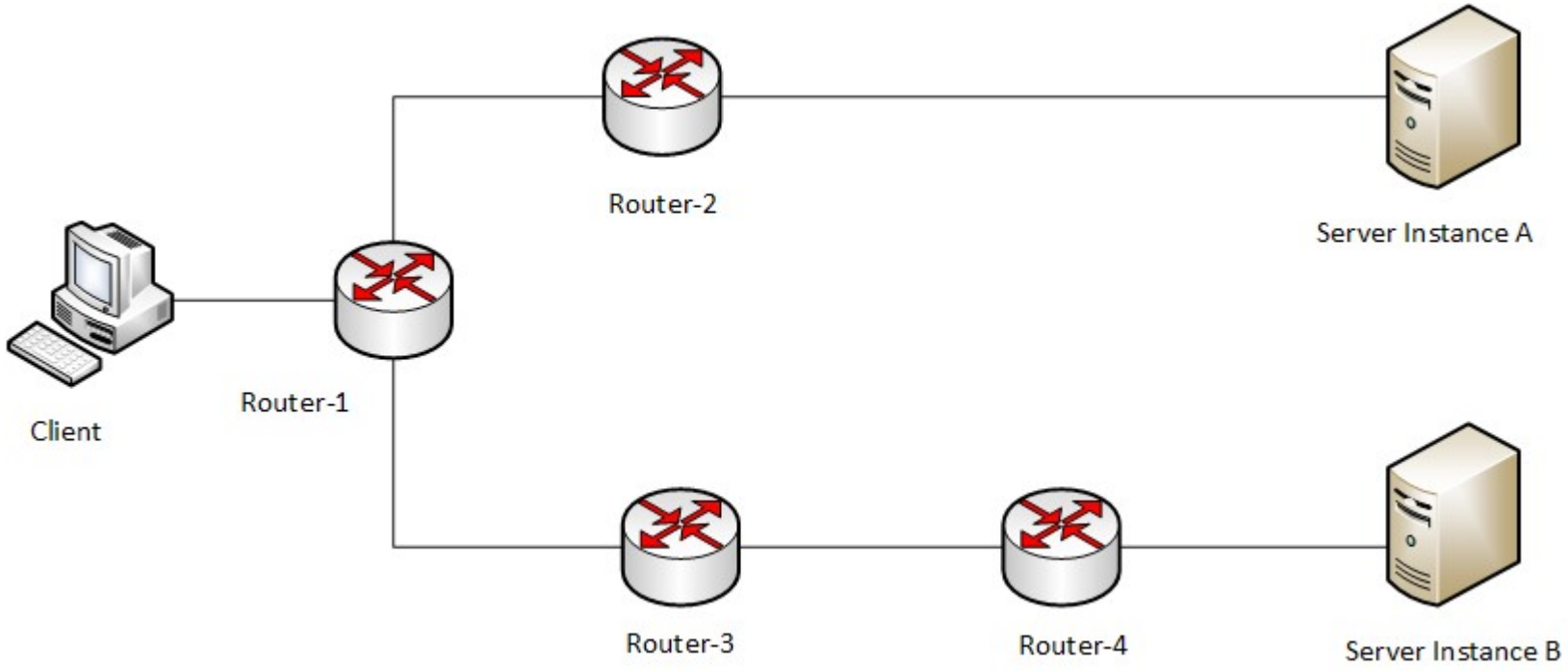
– Multiple instances of a service share the same IP address.

– The routing infrastructure directs any packet to the topologically nearest instance of the service.

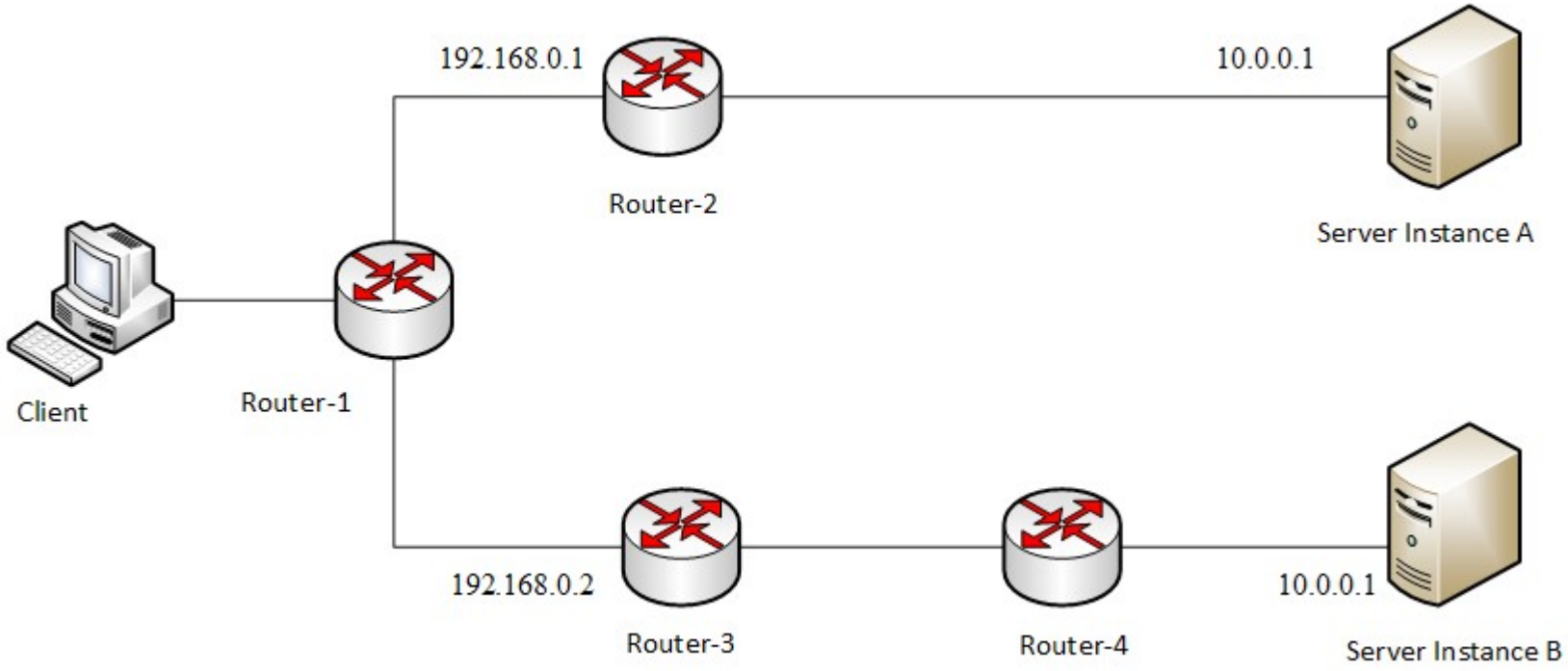
– What little complexity exists is in the optional details.

Lets have an example

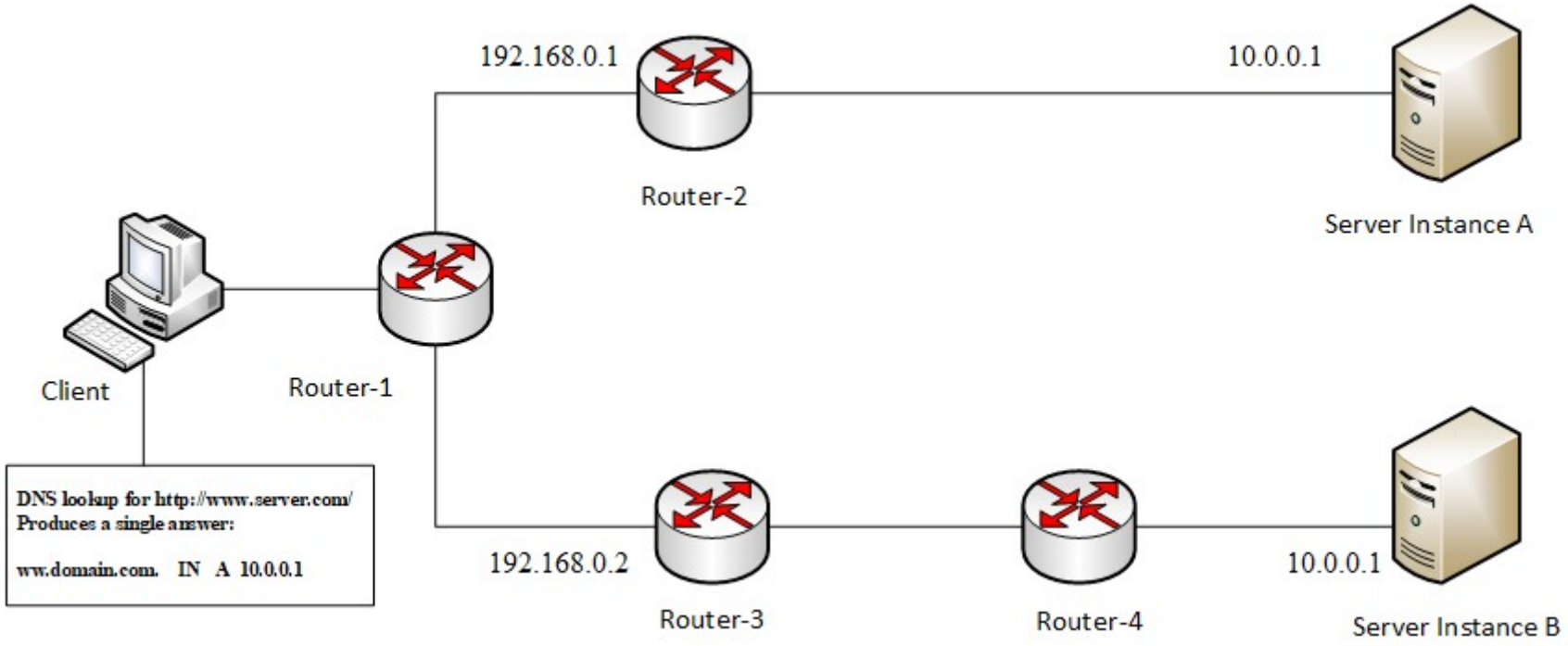
Example – IP Anycast



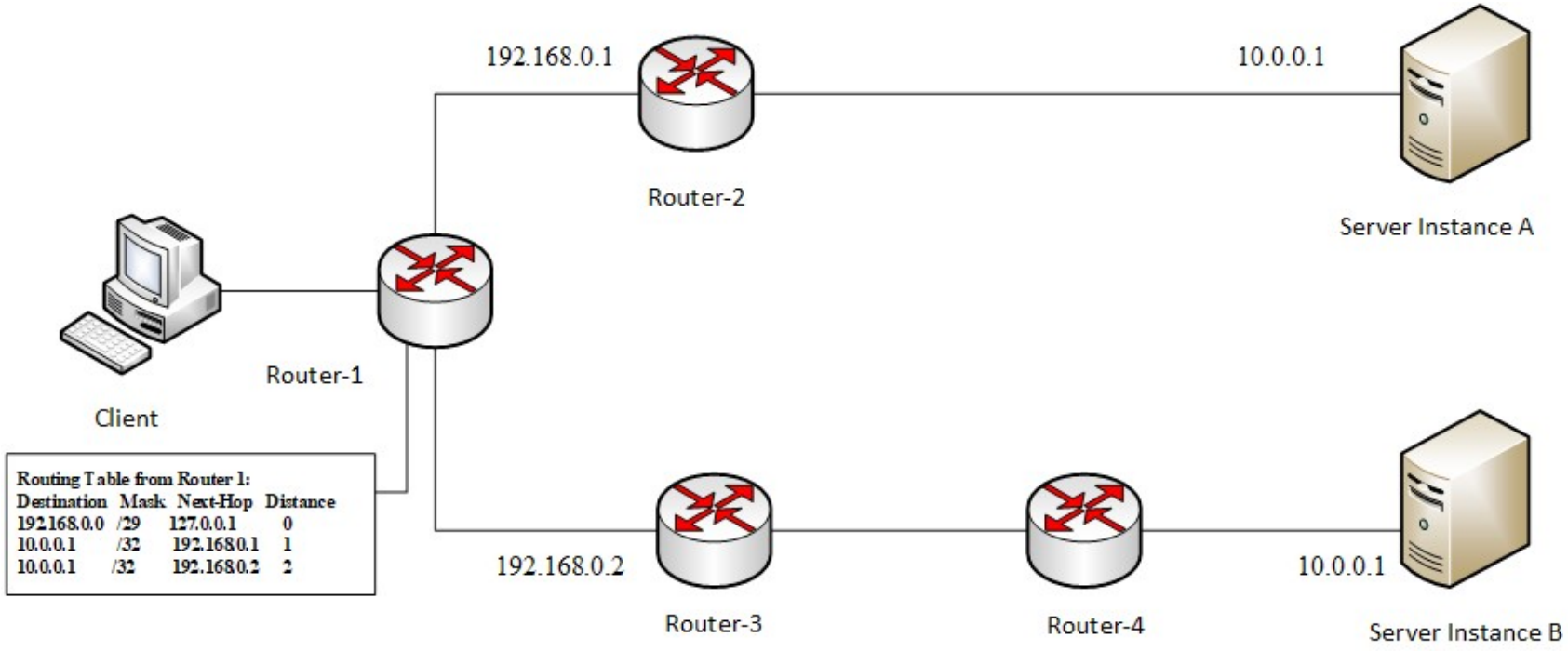
Example – IP Anycast



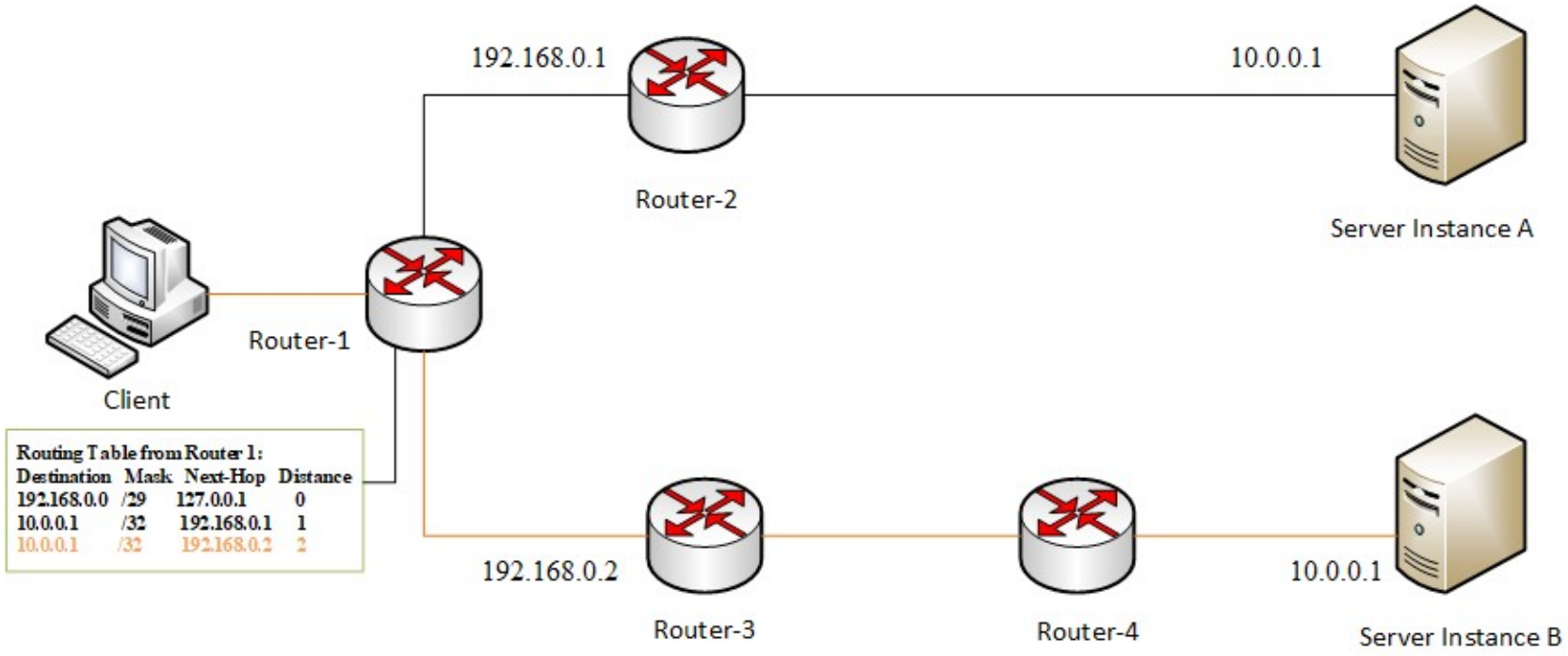
Example – IP Anycast



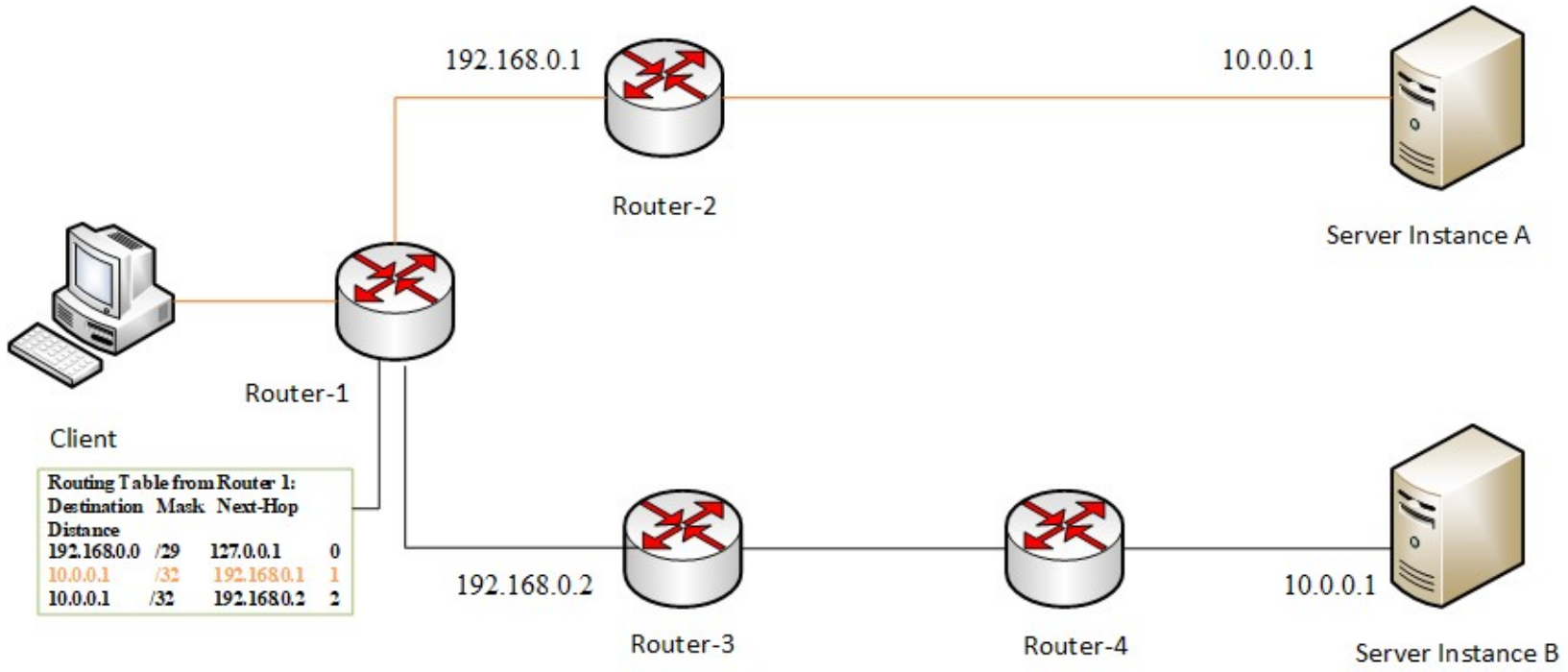
Example – IP Anycast



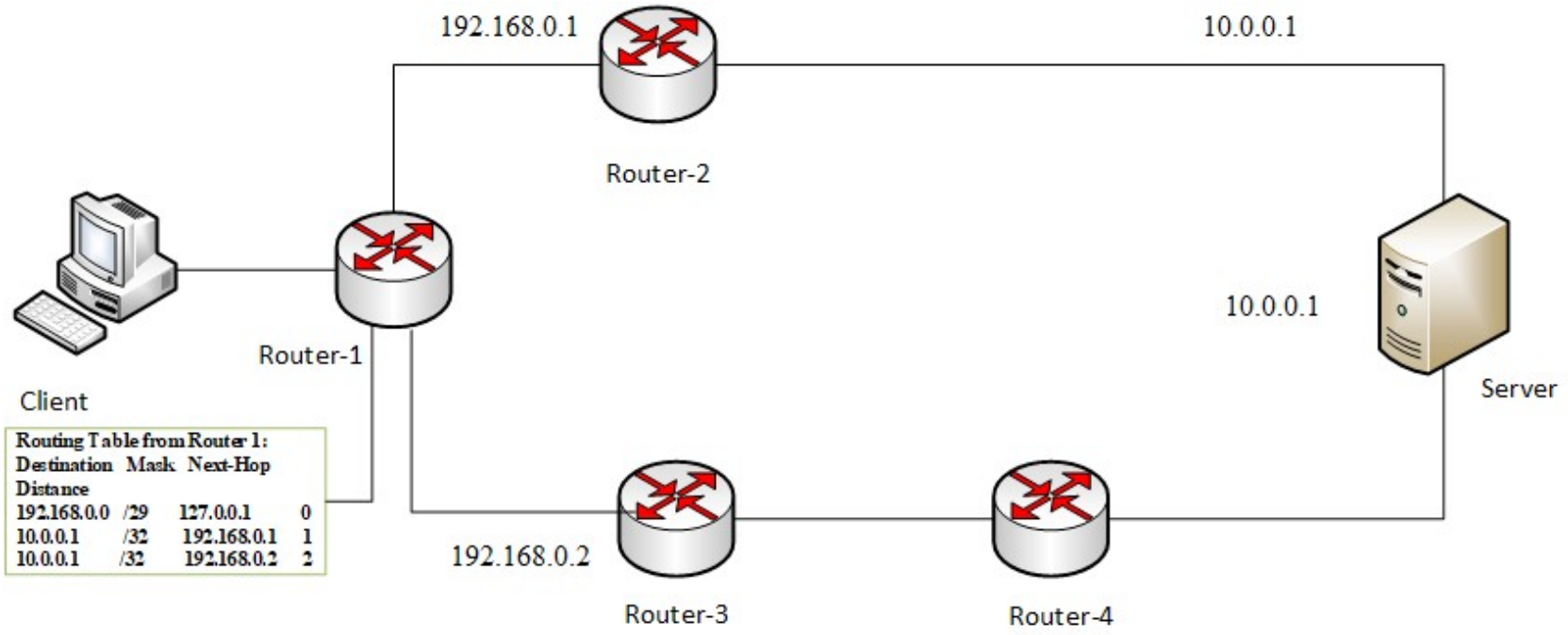
Example – IP Anycast



Example – IP Anycast

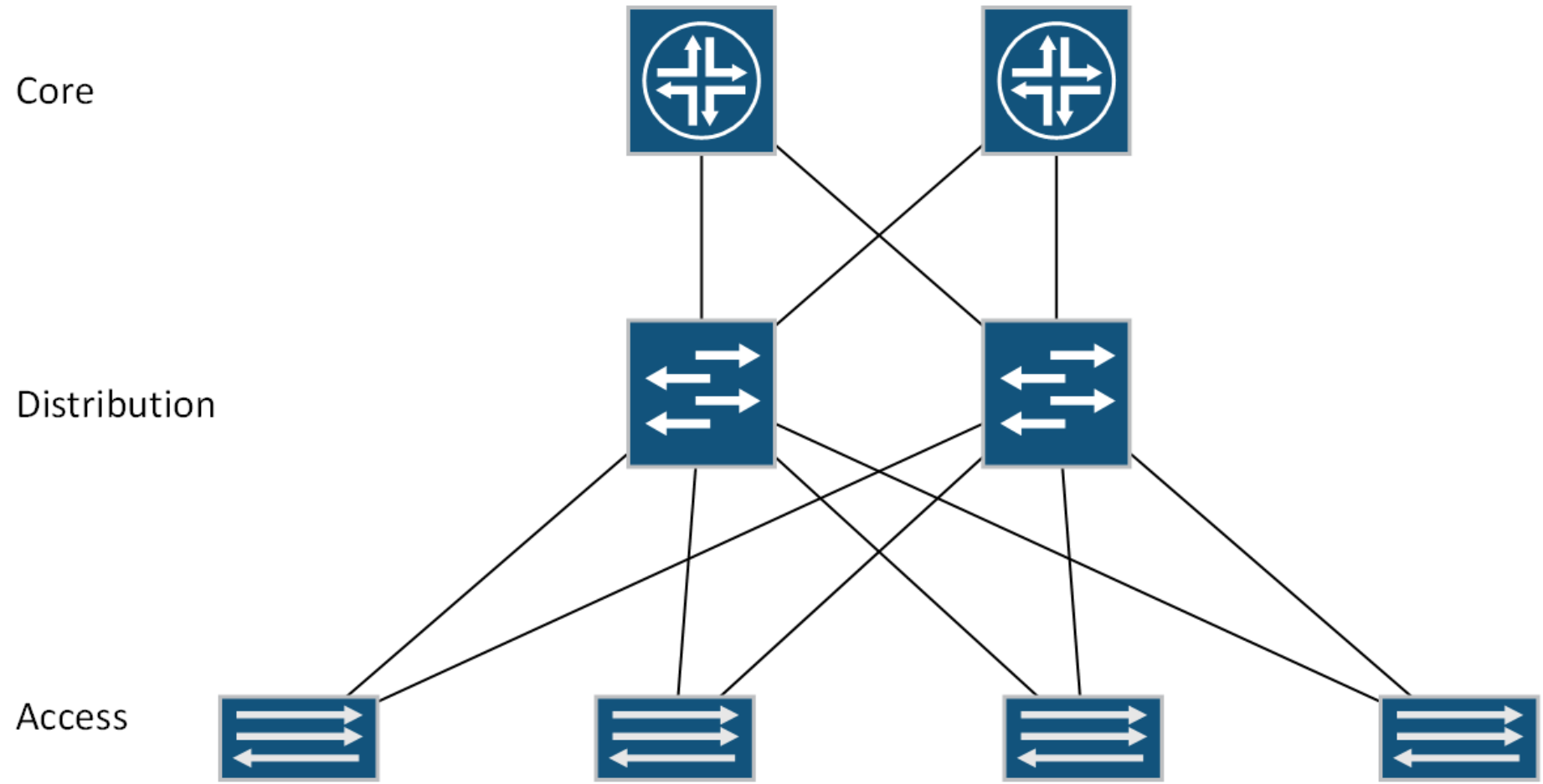


Example – IP Anycast



Network Topology – in consideration

Where to place???



What is DNS ?

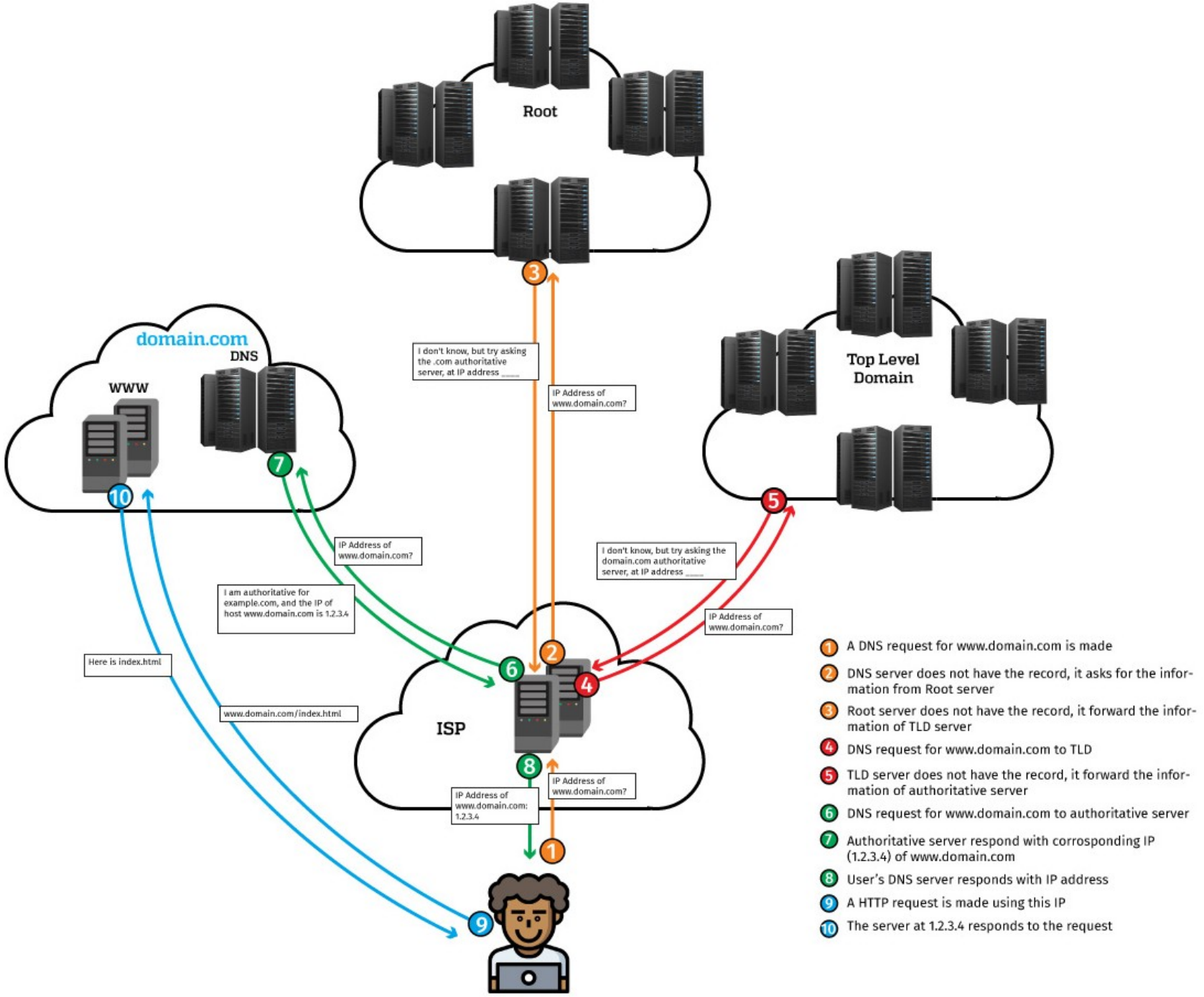
"The Domain Name Server (DNS) is the Achilles heel of the Web. The important thing is that it's managed responsibly."

– Tim Berners-Lee

"DNS is mission-critical to all organizations that connect to the internet. DNS failure or poor performance leads to applications, data and content becoming unavailable, causing user frustration, lost sales and business reputation damage."

– Bob Gill, Gartner, Inc.

How DNS system works?



Recursive DNS – things for consideration

- DNSSEC validation should be enabled
- **chroot** the DNS service
- **allow-query** should be restricted
- **allow-recursion** should be restricted
- Tell everyone about the version, NOT!!!
- **recursive-clients** connection should be mentioned according to the resources
- TCP socket connections should be limited
- **max-cache-size** should be defined, calculating with RAM resources
- Caching controller have to be there
  - mx-cache-ttl, Positive responses should be focused
  - mx-ncache-ttl, Negative responses can not be overlooked
  - min-refresh-time & min-retry-time, also plays vital roll for Caching Recursive DNS system

Anycast Documented

- Concept discussed in RFC 1546
- Operational experience is evolved with the current practices
- Evolution is briefly documented in RFC 2101
- Anycast DNS noted in RFC 2181
- Anycast authoritative name service at RFC3258

DNS Anycast at Link3

Implementation Case Study

Why we choose DNS Anycast – issues we 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 access
4. Log search was not administration friendly
5. No log options for Recursive query

Why we choose DNS Anycast

Because

– 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.

Decision we have taken

– The network topology and location of users should be considered

– The infrastructure restrictions also play a major role

– Main targets high availability and optimal (given the restrictions) load distribution

– Updated service package and OS have to there

– Divide the Authoritative & Recursive in to TWO server

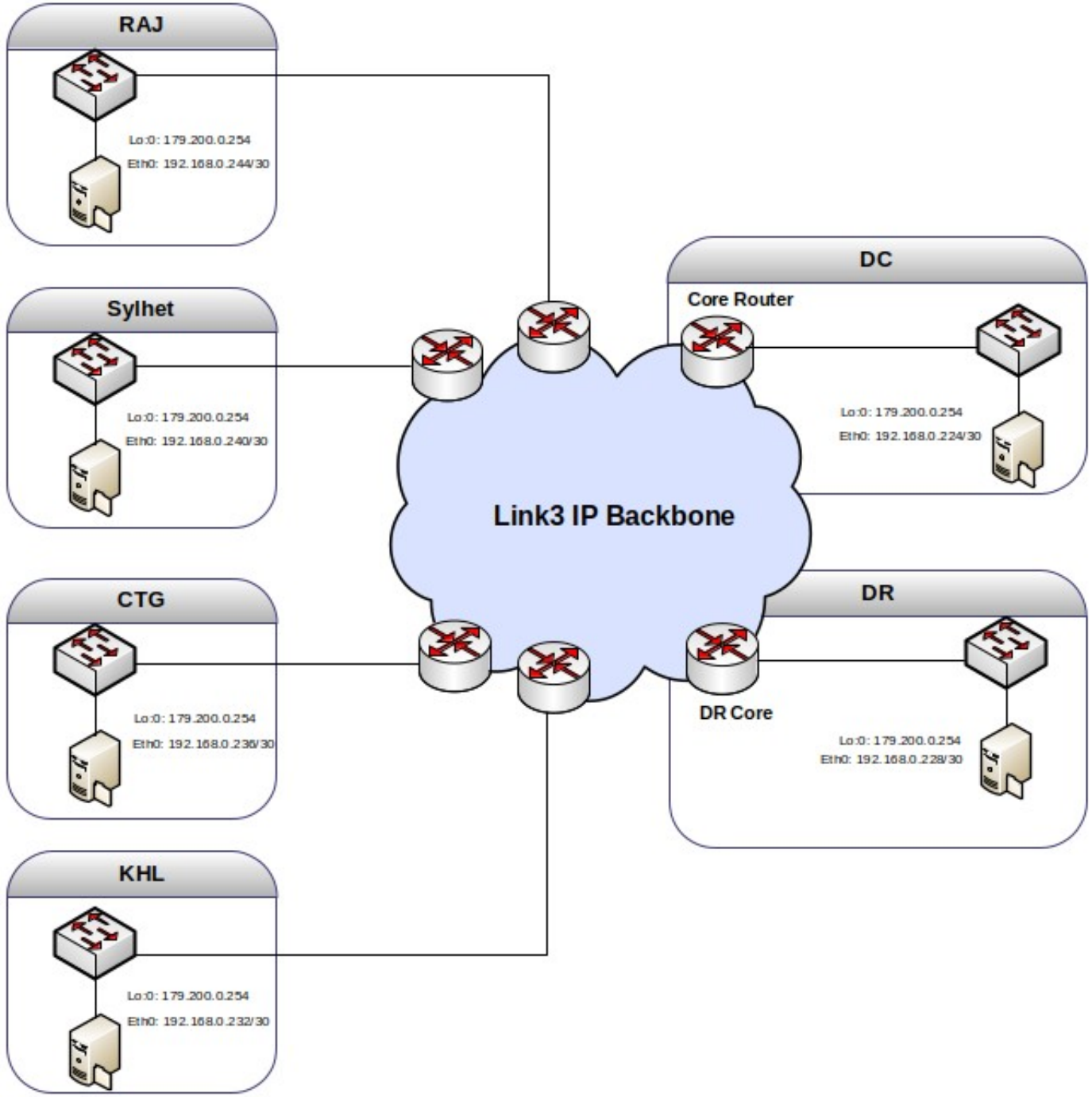
– Deploy the IP Anycast for Recursive DNS only

– Configure the caching log based on search criteria

Resources we have used

Software Resources	Hardware Resources
CentOS 7.5 64 bit	CPU - 4 core 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	

Anycast Infrastructure of Link3



Anycast DNS deployment – can be summarized

- Address selection
- Host configuration
- Service configuration
- Network configuration
- Follow standard security measures
- Check and monitor the service

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

Configuration - host part

– Hosts needed to be configured to accept traffic to anycast address

– A unique management IP address in each host

– Anycast addresses are configured as additional loopbacks

– Filtering for incoming traffic should be updated

Zone-1 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)
```

Zone-1 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;
};
```

Zone-1 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; };
       };
```

Zone-1 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
!
```

Zone-1 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
```

Zone-1 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
```

Zone-2 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)
```

Zone-2 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;
};
```

Zone-2 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; };
       };
```

Zone-2 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
!
```

Zone-2 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
```

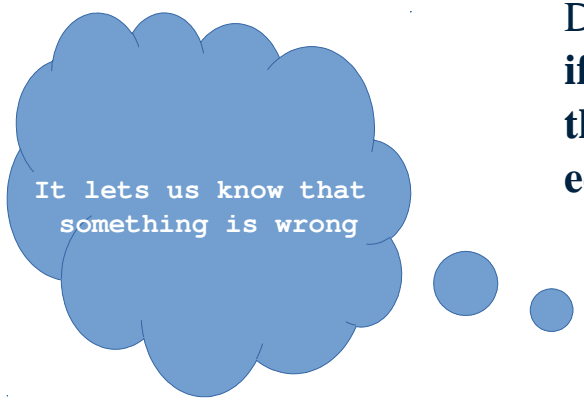
Zone-2 Router - announcing route

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
```

Monitoring the Anycast Service

Anycast is much better when you monitor the service and shutdown the routing announce when the DNS server is down.




```
#!/bin/bash
DNSUP=`/usr/bin/dig @179.100.0.254 localhost. A +short`
if [ "$DNSUP" != "127.0.0.1" ];
then
echo "Stopping Anycast...."
   systemctl bgpd stop
   systemctl zebra stop
     echo "Stopped: Zone-1 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 – Zone-1Anycast DNS
has stooped                                 working" nothing@notification.com
          else
            echo "Everything's good... Do nothing..."
          fi
```

Monitoring of Anycast Services

- Administration of all nodes through central management hosts on trusted LANs
- Each server has IPMI module accessible through the same management LANs
- Monitoring / alerting provided via a Nagios infrastructure already is in place
- DNS measurements from BIND is also added at Nagios NMS.
- PnP4Nagios is in place for graph representation with Nagios

Security Measures

– Hide the bind version

```
[root@bd-anycast-dns etc]# cat /var/named/chroot/etc/named.conf
version “please don’t ask my name" ;
```

– Install & Configure the named service with least privileges CHROOT

```
[root@bd-anycast-dns quagga]# cd /var/named/chroot/ && ls
dev  etc  run  usr  var
```

– 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

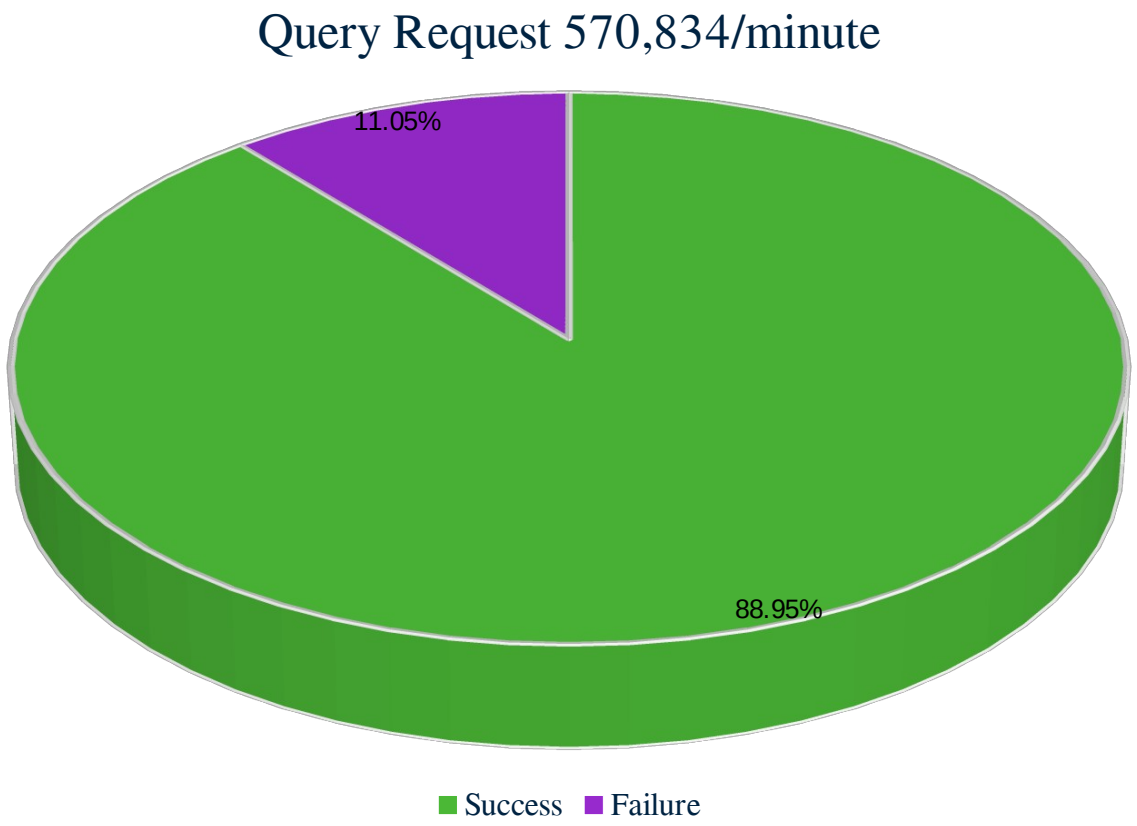
```
[root@bd-anycast-dns etc]# cat /var/named/chroot/etc/named.conf
listen-on port 53 { 127.0.0.1; 179.100.0.254; };
```

Output of Anycast DNS Infrastructure

– serving 822 million queries / Day
– with 5 Caching DNS server

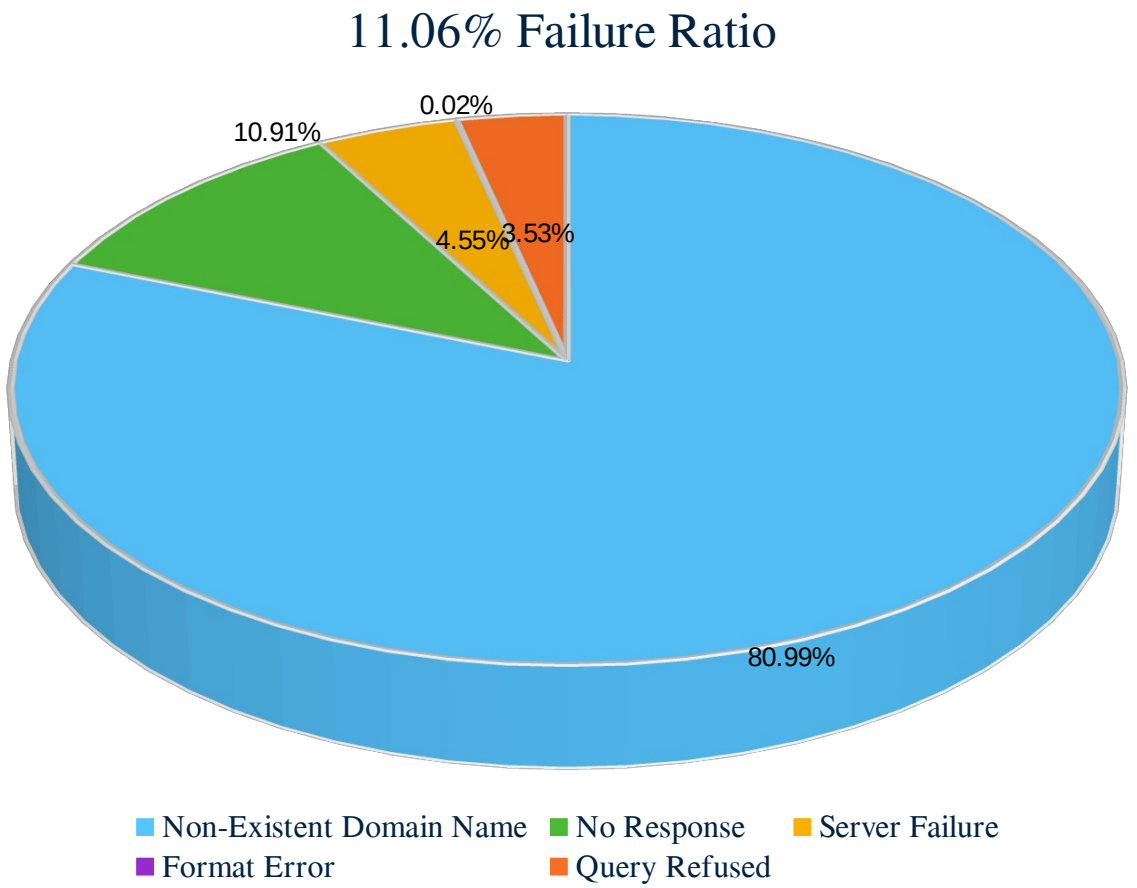
Output of Anycast DNS Infrastructure

Success and Failure ratio



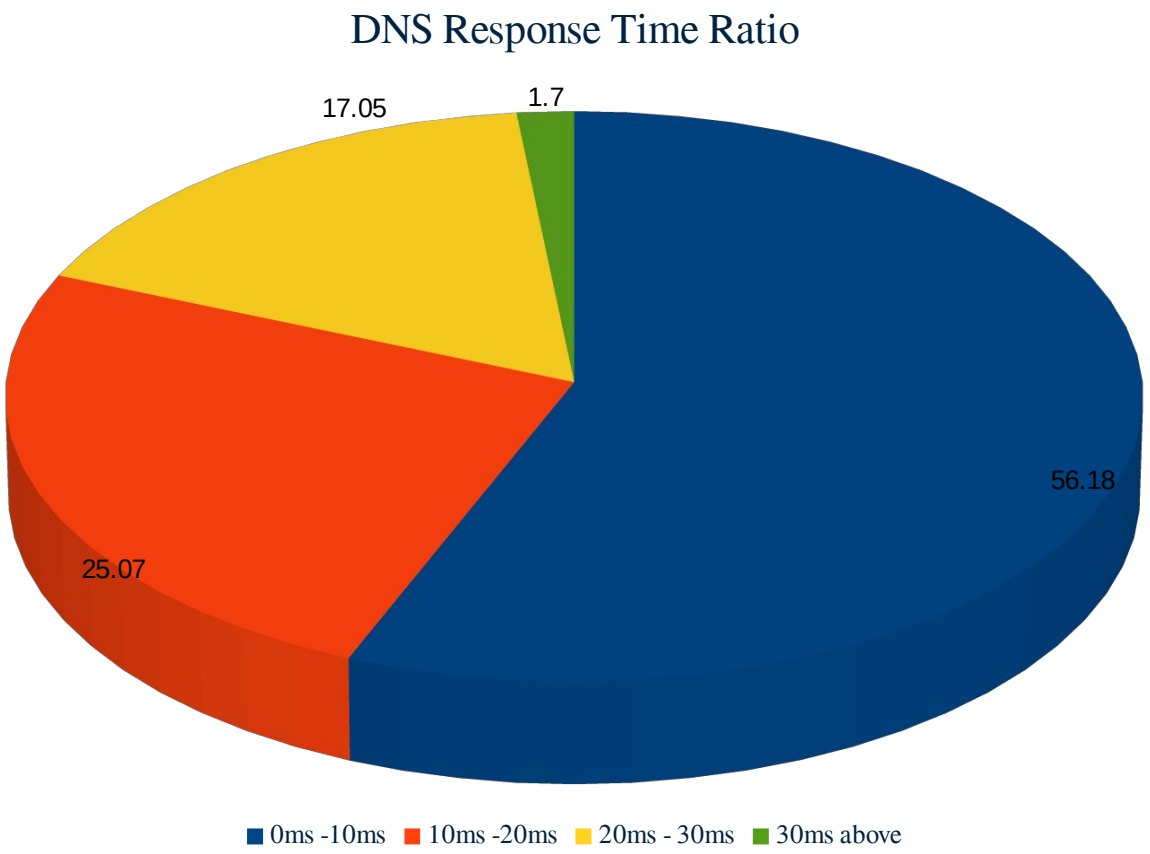
Output of Anycast DNS Infrastructure

DNS Failure Reason



Output of Anycast DNS Infrastructure

DNS Resolution Time



Thank You