Anycast for the DNS

Stéphane Bortzmeyer
AFNIC
bortzmeyer@nic.fr

and Nishal Goburdhan
Afrinic
nishal@afrinic.net
Unicast & Anycast

- **Unicast**: send the message to a specific machine
- **Anycast**: send the message to any of the machines which implement a service (DNS, 6to4...)

In practice, used only when **routing** is used (not load balancers or VRRP). RFC 1546, 4786...
Why anycast?

- Main reason: resilience against denial-of-service attacks. The big accelerator was the great attack against the root in 2002.
- Others reasons: break the size limits of the NS record set.

First DNS deployments: AS 112 (RFC 6304) then the root. Today very common.
General theory of operation

- Several machines listen to the **service IP address**
- Routers announce the service IP address in several places
- Routing algorithm chooses the “closest”

Works with OSPF, BGP or others. On the Internet, we use BGP.
The machines behind a same service IP address are instances of the same anycast cloud.
Let’s see

- **L.root-servers.net** is widely anycasted.
- traceroute from several places to see the different **instances** of L
- All the networks which go to the same instance are an **attraction basin** (watershed?)

Geographical note: the third largest river watershed in the world is in Africa (Congo basin)
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server

- `dig +nsid @l.root-servers.net SOA .`
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server

- `dig +nsid @l.root-servers.net SOA .`
- Unformatted output. `dig` displays as “(h) (e) (r) (0) (1) (.) (l) (.r) (o) (o) (t) (-) (s) (e) (r) (v) (e) (r) (s) (. (o) (r) (g)”
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server

- dig +nsid @l.root-servers.net SOA .
- Unformatted output. dig displays as “(h) (e) (r) (0) (1) (.l) (.r) (o) (o) (t) (-) (s) (e) (r) (v) (e) (r) (s) (.o) (r) (g)”

And from Abidjan?
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server

- `dig +nsid @l.root-servers.net SOA .`
- Unformatted output. dig displays as “(h) (e) (r) (0) (1) (.l) (.r) (o) (o) (t) (-) (s) (e) (r) (v) (e) (r) (s) (.o) (r) (g)”

And from Abidjan? (Old `hostname.bind` not suitable for anycast. Do you see why?)
Name server identity

NSID queries (RFC 5001) allow to know the identity of the name server

- `dig +nsid @l.root-servers.net SOA .`

  Unformatted output. `dig` displays as “(h) (e) (r) (0) (1) (.l) (.r) (o) (o) (t) (-) (s) (e) (r) (v) (e) (r) (s) (.o) (r) (g)”

And from Abidjan? (Old `hostname.bind` not suitable for anycast. Do you see why? See also `http://tools.ietf.org/id/draft-jabley-dnsop-anycast-mapping`
Deploying anycast

**General warning**

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

```bash
DNSISUP=$(dig @$ANYCASTSERVICE $MYDOMAIN SOA +short)
if [ "$DNSISUP" != $GOODANSWER ];
then
  echo "Stopping Anycast...."
  /etc/init.d/bgpd stop
fi
```
Deploying anycast, IGP

(IGP = Internal Gateway Protocol like OSPF)

Deploying anycast, IGP

(IGP = Internal Gateway Protocol like OSPF)

- Useful for recursive name servers (client-based fallback is too slow for a serious service)

Deploying anycast, IGP

(IGP = Internal Gateway Protocol like OSPF)

- Useful for recursive name servers (client-based fallback is too slow for a serious service)
- Also to implement load-sharing on authoritative servers

Deploying anycast, IGP

(IGP = Internal Gateway Protocol like OSPF)

- Useful for recursive name servers (client-based fallback is too slow for a serious service)
- Also to implement load-sharing on authoritative servers

Configuration on the name server

None! Just listen on the service IP address and add the monitoring as seen above.
Configuration on the name server

- None! Just listen on the service IP address
Configuration on the name server

- None! Just listen on the service IP address
- And add the monitoring as seen above
Deploying anycast, EGP

(EGP = External Gateway Protocol, today only BGP)
Deploying anycast, EGP

(EGP = External Gateway Protocol, today only BGP)

- For authoritative name servers
Deploying anycast, EGP

(EGP = External Gateway Protocol, today only BGP)

- For authoritative name servers
- Resiliency is paramount for DNS service (Microsoft’s failure two days ago...)
BGP with Quagga

```text
router bgp 112
bgp router-id x.y.z.t
network 192.175.48.0/24
neighbor a.b.c.d remote-as xxxx
neighbor a.b.c.d prefix-list all in
neighbor a.b.c.d prefix-list as112-out out
```

Yes, that’s all!

https://www.as112.net/as112-centos.html
Other routing software

- For OpenBGPD see https://www.as112.net/as112-freebsd.html
- For BIRD, see http://vincent.bernat.im/en/blog/2011-dns-anycast.html
Origin AS?
Origin AS?

- Two schools of thought: one unique AS (to bind them all :-)

- Some use one unique origin but add an AS per site in the path (you need a lot of AS numbers but Afrinic allows it)
Origin AS?

- Two schools of thought: one unique AS (to bind them all :-)
- Or one AS per instance (RFC 6382)
Origin AS?

- Two schools of thought: one unique AS (to bind them all :-)
- Or one AS per instance (RFC 6382)
- Some use one unique origin but add an AS per site in the path (you need a lot of AS numbers but Afrinic allows it)
Hesitant?

Yes, anycast is not obvious. Start with something less critical: host an instance of AS112!

https://www.as112.net/
Hesitant?

- Yes, anycast is not obvious
Hesitant?

- Yes, anycast is not obvious
- Start with something less critical: host an instance of AS112! [https://www.as112.net/](https://www.as112.net/)
Monitoring anycast

RIPE Atlas probes: a friendly botnet of 4,500 small probes in the world

RIPE stat: a lot of information about routing

http://stat.ripe.net/
Monitoring anycast

- RIPE Atlas probes: a friendly botnet of 4,500 small probes in the world
Monitoring anycast

- RIPE Atlas probes: a friendly botnet of 4 500 small probes in the world
- RIPE stat: a lot of information about routing
  http://stat.ripe.net/
Buying anycast

There are also several providers who lease you anycast hosting services.
Buying anycast

There are also several providers who lease you anycast hosting services. You still have to monitor and to check (in one african country, the provider claimed they have an anycast instance in the country, which was false)
Catching fire

(Thanks to S. Collins)
Catching fire

(Thanks to S. Collins)

- Better resilience against dDoS
Catching fire

(Thanks to S. Collins)

- Better resilience against dDoS
- Attack is contained in one attraction basin (local attacks stay local)
Conclusion

A technology now mature
Which seriously improves DNS quality and resiliency

20 / 20
Conclusion

- A technology now mature
Conclusion

- A technology now mature
- Which seriously improves DNS quality and resiliency
Merci !