Anycast vs. DDoS: Evaluating Nov. 2015 Root DNS event

Giovane C. M. Moura¹, Ricardo de O. Schmidt², John Heidemann³, Wouter B. de Vries², Moritz Müller¹, Lan Wei³, Cristian Hesselman¹

¹SIDN Labs ²University of Twente ³USC/ISI

2016-07-18
IEFT 96 – MAPRG/IRTF
IP Anycast

* It's simple: “making a particular Service Address available in multiple, discrete, autonomous locations” (RFC4786, 7094)

* Improves performance and resilience (1 IP → Many services, 1 down, others operate)

* Widely use in DNS (and also CDNs)
DDoS

* Getting bigger (400Gbps +)

* Getting cheaper (booters, few dollars)

* Happening more often

* Core idea: bring down services

* Question: *How anycast behaves during a DDoS attack?*

* Case study: Root DNS events Nov 2015
The Root DNS system

- List the records that points to all TLDs (.com, .nl, .net...)

<table>
<thead>
<tr>
<th>letter</th>
<th>operator</th>
<th>sites (global, local)</th>
<th>architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Verisign</td>
<td>5 (5, 0)</td>
<td>anycast</td>
</tr>
<tr>
<td>B</td>
<td>USC/ISI</td>
<td>1 (1, 0)</td>
<td>single site</td>
</tr>
<tr>
<td>C</td>
<td>Cogent</td>
<td>8 (8, 0)</td>
<td>anycast</td>
</tr>
<tr>
<td>D</td>
<td>U. Maryland</td>
<td>87 (18, 69)</td>
<td>anycast</td>
</tr>
<tr>
<td>E</td>
<td>NASA</td>
<td>12 (1, 11)</td>
<td>anycast</td>
</tr>
<tr>
<td>F</td>
<td>ISC</td>
<td>59 (5, 54)</td>
<td>anycast</td>
</tr>
<tr>
<td>G</td>
<td>U.S. DoD</td>
<td>6 (6, 0)</td>
<td>anycast</td>
</tr>
<tr>
<td>H</td>
<td>ARL</td>
<td>2 (2, 0)</td>
<td>primary/backup</td>
</tr>
<tr>
<td>I</td>
<td>Netnod</td>
<td>49 (49, 0)</td>
<td>anycast</td>
</tr>
<tr>
<td>J</td>
<td>Verisign</td>
<td>98 (66, 32)</td>
<td>anycast</td>
</tr>
<tr>
<td>K</td>
<td>RIPE</td>
<td>33 (15, 18)</td>
<td>anycast</td>
</tr>
<tr>
<td>L</td>
<td>ICANN</td>
<td>144 (144, 0)</td>
<td>anycast</td>
</tr>
<tr>
<td>M</td>
<td>WIDE</td>
<td>7 (6, 1)</td>
<td>anycast</td>
</tr>
</tbody>
</table>

Table: The 13 Root Letters, each operating a separate DNS service, and their number of sites and architecture as of 2015-11-18.
A Bad Day at the Root...

Data: RIPE DNSmon
Red: >30% loss
(some sites ~99% loss!)

What happened?
What does “red” really mean?
Anycast vs. DDoS in general?
Summary of the Events

Two events

- 2015-11-30t06:50 for 2h40m
- 2015-12-01t05:10 for 1h

affected 10 of 13 letters

about 5M q/s or 3.5Gb/s per affected letter

- aggregate: 155Gb/s

real DNS queries, common query names, from spoofed source IPs

implications:

some letters had high loss

overall, though DNS worked fine

- clients retried other letters (as designed)

data:
A-Root had full view (Verisign presentation);
RSSAC-002 reports
How Well Does Anycast Defend?

561 root DNS locations for 13 services (in 2016-01)

is 561 too few? too many?

what happens under stress?
Anycast in Good Times

Anycast matches a user to a (hopefully) nearby site.

Anycast divides the Internet into catchment (often messy and non-geographic).

(some sites have more capacity)
Anycast Under Stress

too many attackers overwhelm your site: your queries get lost

A similar size attack may be absorbed at a bigger site

Catchments also isolate sites from attackers
Anycast Under Stress (do nothing)

1. **nothing**: X-SJC is **degraded absorber**, protecting X-SYD’s users
Anycast Under Stress (withdraw some routes)

1. **nothing**: X-SJC is degraded absorber, protecting X-SYD’s users.

2. **withdraw** routes from X-SJC; may shift attackers to big site.
Anycast Under Stress (withdraw other routes)

1. **nothing**: X-SJC is degraded absorber, protecting X-SYD’s users.
2. **withdraw** routes from X-SJC; may shift attackers to big site.
3. **withdraw** wrong routes from X-SJC; may shift attackers to other site.

- attackers
- your friend
- you
- X-SJC
- X-PRG
- X-SYD
- another friend
Best reaction to stress: you don't know

1. nothing: X-SJC is degraded absorber, protecting X-SYD’s users
2. withdraw routes from X-SJC; may shift attackers to big site
3. withdraw wrong routes from X-SJC; may shift attackers to other site

Don’t fully control routing and catchments

Don’t know:
- number of attackers
- location of attackers
- affects of routing change

Hard to make informed choices
What Actually Happens?

studying Nov. 30
we see withdrawals and degraded absorbers
some clients loose service
• results vary by anycast deployment
Data About Nov. 30

RIPE Atlas
• 9000 vantage points (RIPE Atlas probes)
• try every letter every 4 minutes
  • except A-root, at this time, was every 30 minutes
• data-plane queries
• global, but heavily biased to Europe

RSSAC-002 reports
• self-reports from letters
• not guaranteed when under stress

BGPmon routing
• control plane
How About the Letters?

some did great:  
D, L, M: not attacked  
A: no visible loss

most suffered:  
a bit (E, F, I, J, K)  
or a lot (B, C, G, H)

but does “x%” measure what users actually see?
Reachability at K-sites

sites see fewer VPs, but why?
- query loss? site absorbs attack,
  but sad customers
Site Flips from Routing Changes

360 minutes (in 4 minute bins)

Nov. 30 event

stay at K-LHR;
sad during event

flip to K-AMS;
(less) sad during event;
back to K-LHR after flip to K-other
and stay there
flip to K-AMS

yellow: K-LHR
blue: K-AMS
white: K-other
black: failed query

[Moura16a, figure 11b; data: RIPE Atlas]
Confirming flips in BGP

flips common during events for most letters

flips seen in BGP
Flips Across Letters: E and K

to evaluate flips over two days:
compare minimum and maximum cachement
(measured in VPs/site)

(normalize to median
(natural catchment)

(sites acquiring
VPs (during event?)

(sites shedding
VPs

(red sites: <20 VPs; not enough to provide meaningful results)

[Moura 06, figure 5: data: RIPE Atlas]

USC University of Southern California
UNIVERSITY OF TWENTE.
SIDN LABS
Flips Implications

some ISPs are “sticky” and won’t flip
  • will suffer if their site is overloaded
some ISPs will flip
  • but new site may not be much better

result depends on many factors
  • actions taken by root operator
  • routing choices by operator and peer
    • and perhaps peer’s peers, depending on congestion location
      implementation choices
  • DNS, routing
Aside: Collateral Damage

can an event hurt non-targets?

yes! ...a risk of shared datacenters

D-FRA and D-SYD: less traffic
(even though D was not directly attacked)

.NL-FRA and .NL-AMS: no traffic

In other attacks, B-Root’s ISP saw loss to other customers
Conclusions

anycast under stress is complicated

- some users will see persistent loss
- “x% loss” is not complete picture

reactions depend on design and implementation choices

- many not under operator control

more info:

paper: [http://www.isi.edu/~johnh/PAPERS/Moura16a/](http://www.isi.edu/~johnh/PAPERS/Moura16a/)
data: [https://ant.isi.edu/anycast/](https://ant.isi.edu/anycast/)