

draft-moura-dnsop-authoritative-recommendations-03

Giovane C. M. Moura^{1,2}, Wes Hardaker³,
John Heidemann³, Marco Davids¹

DNSOP – IETF 104
Prague, CZ
2019-03-XX

¹SIDN Labs, ²TU Delft, ³USC/ISI

Draft History

- This is an **Informational** draft
- **Today:** first time presented at DNSOP
- Versions and mailing list discussion:
 - **-03 (2019-03-11):** (minor changes from -02)
 - **-02 (2019-03-08):** [link list thread \(no responses\)](#)
 - **-01 (2018-12-20):** [link list thread \(no responses\)](#)
 - **-00 (2018-11-28):** [link list thread](#)
- Github link:
 - [https://github.com/gmmoura/
draft-moura-dnsop-authoritative-recommendations](https://github.com/gmmoura/draft-moura-dnsop-authoritative-recommendations)

- 13 people that have had 5 research papers:
 - Draft authors + Ricardo de O Schmidt, Wouter B. de Vries, Moritz Müller, Lan Wei, Cristian Hesselman, Jan Harm Kuipers, Pieter-Tjerk de Boer and Aiko Pras.
- Relevant papers with *recommendations* backed by large-scale, Internet-wide measurements:
 - 4x ACM IMC
 - 1x PAM
- However, papers tend to be *long, detailed* – they explain *why*
- **Target group:** large authoritative DNS ops, with global traffic

This draft:

```
papers = []
papers.append(Moura16b)
papers.append(Mueller17b)
papers.append(Schmidt17a)
papers.append(Vries17b)
papers.append(Moura18b)

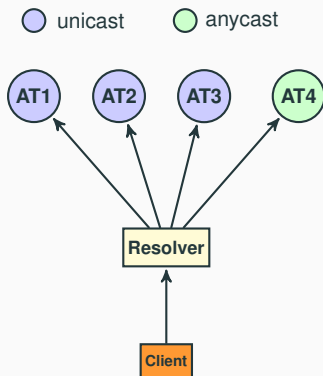
for p in papers:
    recommendations = TLDR(p) #great filter :-)
    print(recommendations)
```

- Tangible, direct language to OPs folks on *what* to
- Reader referred to papers to know *why*

Recommendations in a nutshell

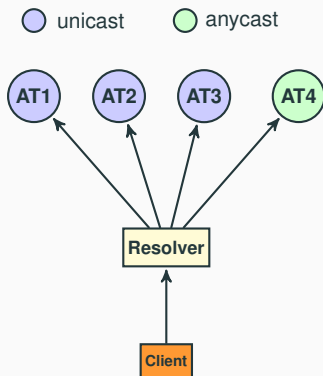
- R1: Use equally strong IP anycast in every authoritative server to achieve even load distribution [1]
- R2: Routing Can Matter More Than Locations [2]
- R3: Collecting Detailed Anycast Catchment Maps Ahead of Actual Deployment Can Improve Engineering Designs [3]
- R4: When under stress, employ two strategies [4]
- R5: Consider longer time-to-live values whenever possible [5]
- R6: Shared Infrastructure Risks Collateral Damage During Attacks [4]

R1: Use equally strong IP anycast in every authoritative server to achieve even load distribution



- **Resolvers** will query **ALL** authoritatives [1]
 - Ripe Atlas, **.nl** and the Roots data
- However, ATs nearby will get more queries

R1: Use equally strong IP anycast in every authoritative server to achieve even load distribution



- For OPs: latency of *all* ATs matter
- Unicast, by definition, cannot deliver good global performance
- [1] recommends then use anycast in *all* NS records
 - equally strong (peering and capacity), and phase out unicast.
- This has been applied to **.nl**.

R2: Routing Can Matter More Than Locations

- When choosing an anycast DNS provider, people always ask “how many sites/instances” it has
- Assumption: more sites \rightarrow lower latency
- [2] shows that this is not always true:
 - c-root: 8 locations.
 - k-root: 33 locations
 - l-root: 144 locations
 - Their **median RTT: 30–32 ms** to 7.9k Atlas probes

R2: Routing Can Matter More Than Locations

- Why? BGP is agnostic to geographical distance
 - A California client may be answered by a site in NRT – even though there is a site in SFO
- [2] thus recommends to consider **routing** and **connectivity** when engineering DNS anycast services
 - 12 sites is enough to provide good global latency
 - However, more sites may be helpful in case of DDoS [4]

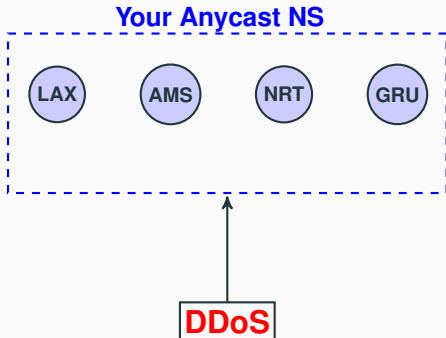
R3: Collecting Detailed Anycast Catchment Maps Ahead of Actual Deployment Can Improve Engineering Designs

- Say you run an anycast service with n instances
- Say you want to add 1 more instance in SFO
- How will that affect traffic among your other locations?
 - **Very hard** to predict

R3: Collecting Detailed Anycast Catchment Maps Ahead of Actual Deployment Can Improve Engineering Designs

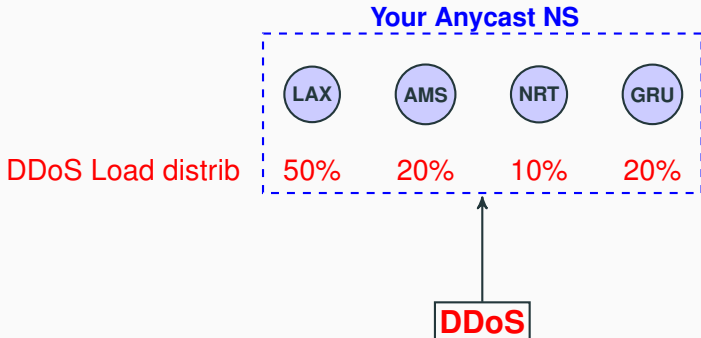
- Solution: detailed anycast catchment maps
- [3] present a tool (Verfploeter) that does that using ICMP
- They predicted b-root catchments and query loads:
 - Load predict going to b-root LAX instance: 81.6%
 - Actual load: 81.4%.
- Current deployments:
 1. Anycast testbed (<http://anycast-testbed.nl>)
 2. B-root
 3. Large unnamed operator

R4: When under stress, employ two strategies



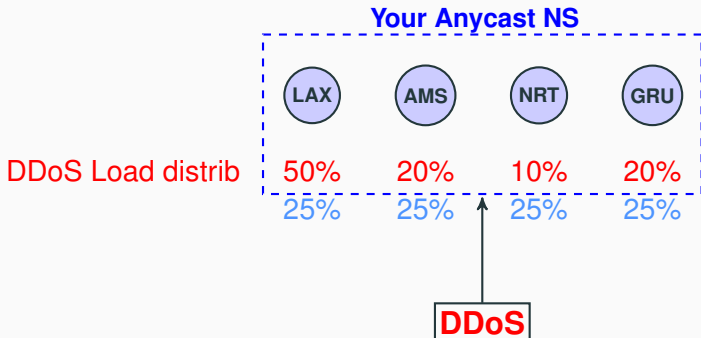
- BGP will map traffic to locations
- What to do? Depends on the attack
 1. **Do nothing and let LAX become a degraded absorber**
 2. **Withdraw/prepend routes to shift traffic**
- Best option depends on attack and NS specifics

R4: When under stress, employ two strategies



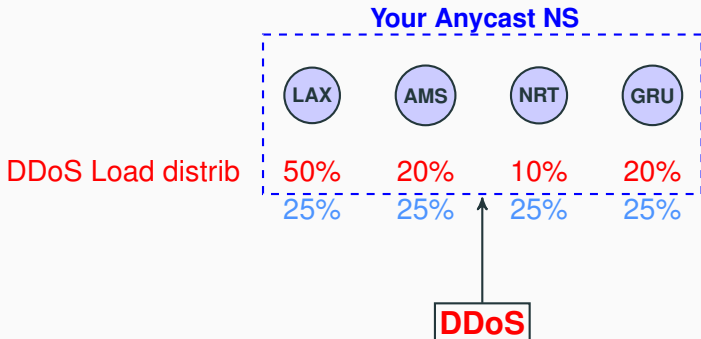
- BGP will map traffic to locations
- What to do? Depends on the attack
 1. **Do nothing and let LAX become a degraded absorber**
 2. **Withdraw/prepend routes to shift traffic**
- Best option depends on attack and NS specifics

R4: When under stress, employ two strategies



- BGP will map traffic to locations
- What to do? Depends on the attack
 1. **Do nothing and let LAX become a degraded absorber**
 2. **Withdraw/prepend routes to shift traffic**
- Best option depends on attack and NS specifics

R4: When under stress, employ two strategies



- BGP will map traffic to locations
- What to do? Depends on the attack
 1. **Do nothing and let LAX become a degraded absorber**
 2. **Withdraw/prepend routes to shift traffic**
- Best option depends on attack and NS specifics

R5: Consider longer TTL values whenever possible

- TTLs set how long queries should remain in resolver's cache
 - Sort of “ephemeral replication”
- They also emulate DDoS attacks (50-100% packet loss)

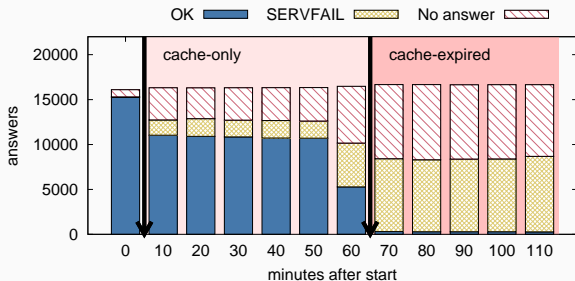


Figure 1: TTL: 1h; 100% Packet loss after $t = 10\text{min}$

R5: Consider longer TTL values whenever possible

- Caching is a *key* component of resolver's resilience
- Retries as well – to the point that resolvers may “hammer” authoritatives
- As such, [5] recommend longer TTLs whenever possible

R6: Shared Infrastructure Risks Collateral Damage During Attacks

- Be careful when hiring/engineering DNS services: co-location implies you shared some (parts of the) infrastructure
- [4] found that co-located `.nl` sites suffered during DDoS against Roots
- Dyn 2016 Attack shows the same
- OPS: be aware of shared infrastructure risk

Questions?

- R1: Use equally strong IP anycast in every authoritative server to achieve even load distribution [1]
- R2: Routing Can Matter More Than Locations [2]
- R3: Collecting Detailed Anycast Catchment Maps Ahead of Actual Deployment Can Improve Engineering Designs [3]
- R4: When under stress, employ two strategies [4]
- R5: Consider longer time-to-live values whenever possible [5]
- R6: Shared Infrastructure Risks Collateral Damage During Attacks [4]

References I

- [1] M. Müller, G. C. M. Moura, R. de O. Schmidt, and J. Heidemann, “Recursives in the wild: Engineering authoritative DNS servers,” in *Proceedings of the ACM Internet Measurement Conference*, London, UK, 2017, pp. 489–495. [Online]. Available: <http://www.isi.edu/%7ejohnh/PAPERS/Mueller17b.html>
- [2] R. d. O. Schmidt, J. Heidemann, and J. H. Kuipers, “Anycast latency: How many sites are enough?” in *Proceedings of the Passive and Active Measurement Workshop*. Sydney, Australia: Springer, Mar. 2017, p. to appear, awarded Best Paper. [Online]. Available: <http://www.isi.edu/%7ejohnh/PAPERS/Schmidt17a.html>

- [3] W. B. de Vries, R. de O. Schmidt, W. Hardaker, J. Heidemann, P.-T. de Boer, and A. Pras, “Verfploeter: Broad and load-aware anycast mapping,” in *Proceedings of the ACM Internet Measurement Conference*, London, UK, 2017. [Online]. Available: <http://www.isi.edu/%7ejohnh/PAPERS/Vries17b.html>
- [4] G. C. M. Moura, R. de O. Schmidt, J. Heidemann, W. B. de Vries, M. Müller, L. Wei, and C. Hesselman, “Anycast vs. DDoS: Evaluating the November 2015 root DNS event,” Nov. 2016. [Online]. Available: <https://www.isi.edu/%7ejohnh/PAPERS/Moura16b.html>

- [5] G. C. M. Moura, J. Heidemann, M. Müller, R. de O. Schmidt, and M. Davids, “When the dike breaks: Dissecting DNS defenses during DDoS,” in *Proceedings of the ACM Internet Measurement Conference*, Oct. 2018. [Online]. Available: <https://www.isi.edu/%7ejohnh/PAPERS/Moura18b.html>