# TsuKing: Coordinating DNS Resolvers and Queries into Potent DoS Amplifiers

Wei Xu, Xiang Li, **Chaoyi Lu**, Baojun Liu, Jia Zhang,
Jianjun Chen, Haixin Duan, Tao Wan

*Tsinghua University; CableLabs*

November 2023

# Name of Attack - Breakdown

**Tsu-King ?**

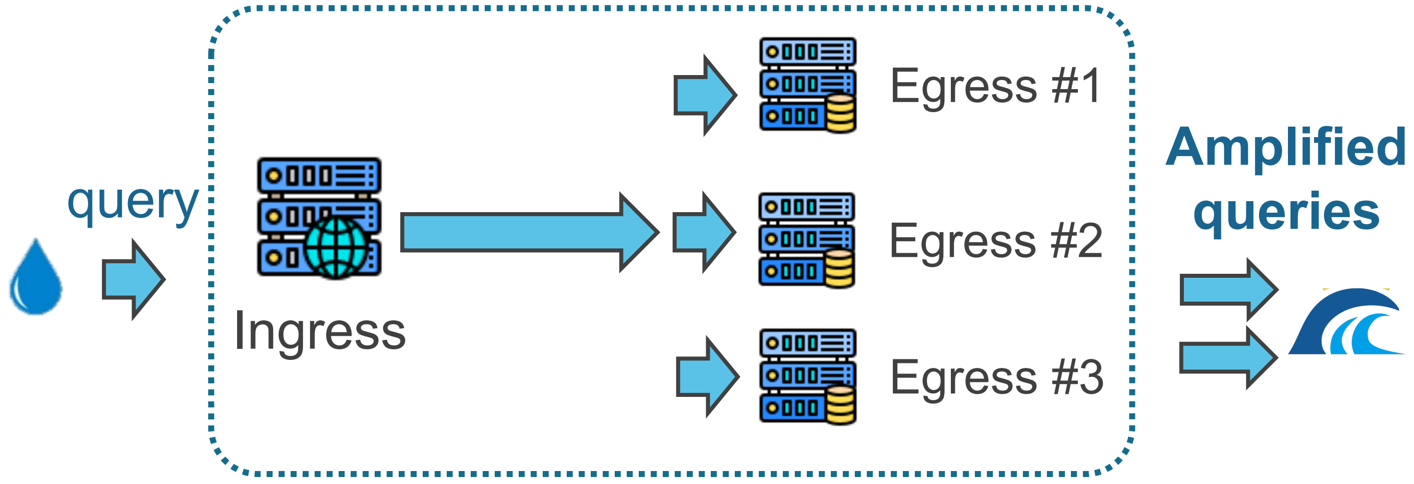
# Name of Attack - Breakdown

## Tsu-King

## Tsunami

*(Traffic amplification ability)*

- Cause: DNS implementation choices & complex infrastructure



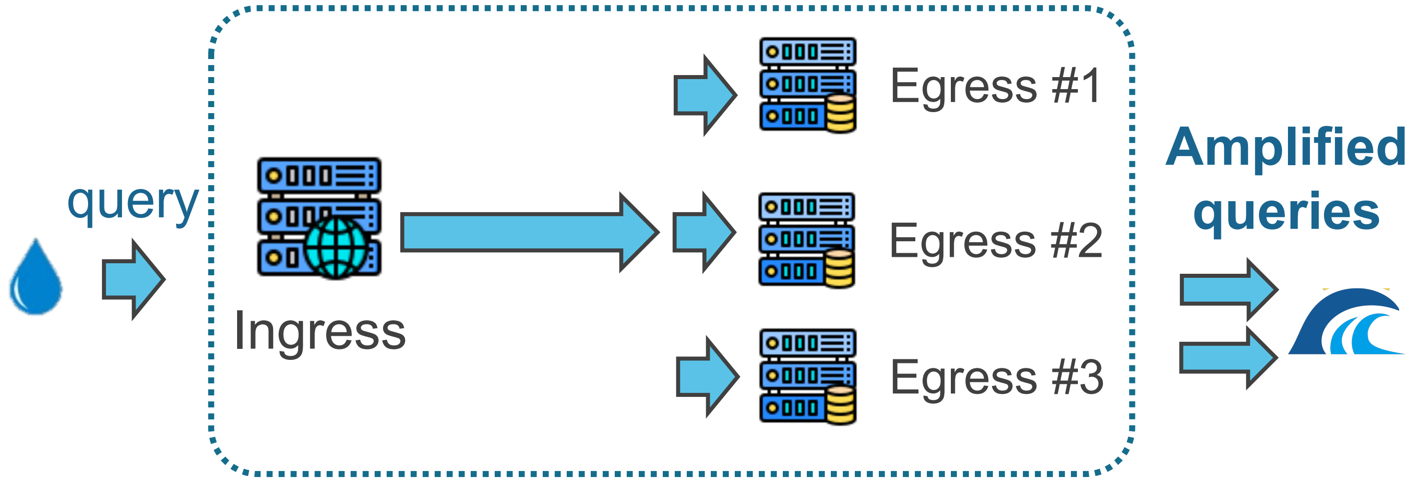
# Name of Attack - Breakdown

## Tsu-King

### Tsunami

*(Traffic amplification ability)*

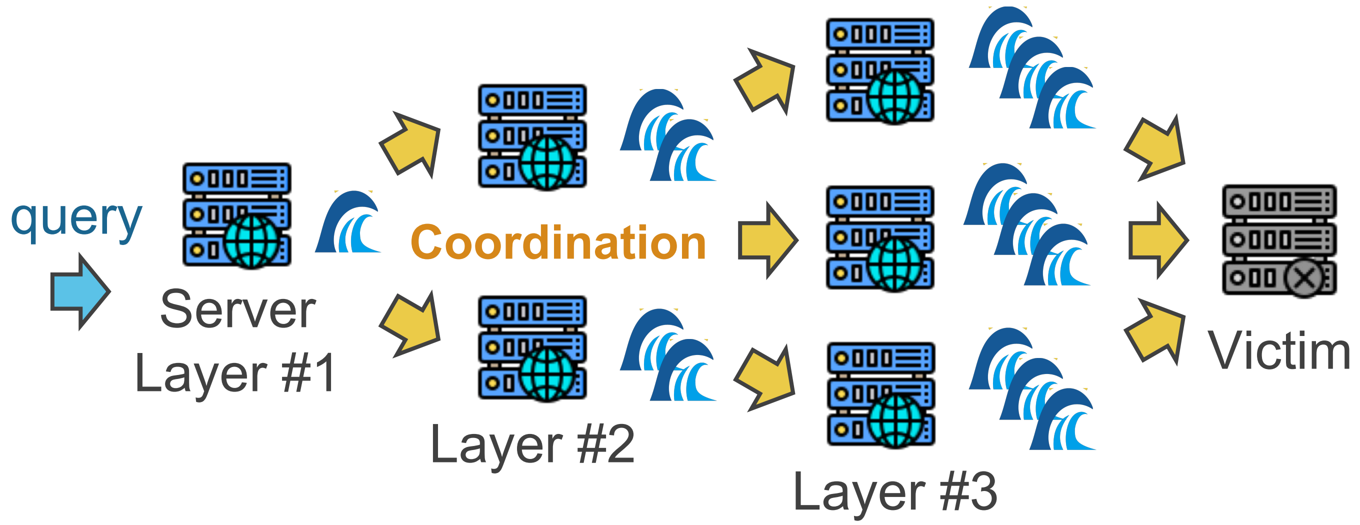
- Cause: DNS implementation choices & complex infrastructure




### King

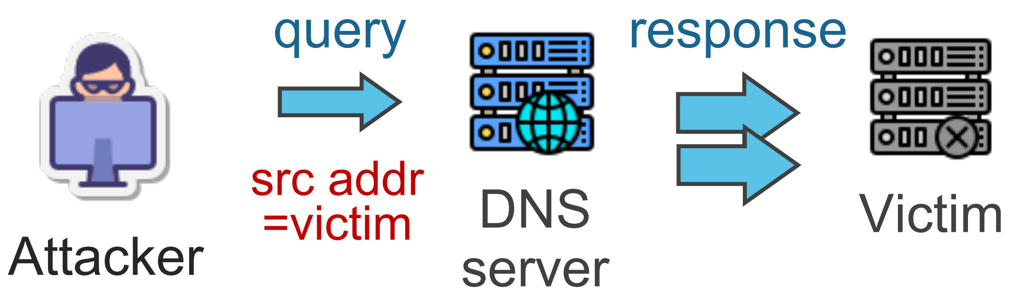
*(Server coordination ability)*

- Coordinates DNS server systems -> **3,000+× amplification factor (*king* of DoS)**



# DNS as a classic DoS attack vector

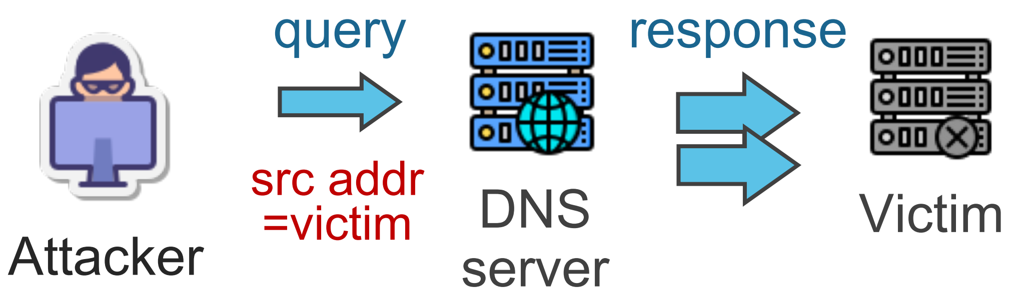
- **Design choices of the DNS protocol**
  - Runs over UDP ➔ *reflected DoS attacks possible*
  - Response larger than query ➔ *traffic amplification*




Reflected DoS attack via DNS

# DNS as a classic DoS attack vector

- **Design choices of the DNS protocol**
  - Runs over UDP ➔ *reflected DoS attacks possible*
  - Response larger than query ➔ *traffic amplification*




Reflected DoS attack via DNS

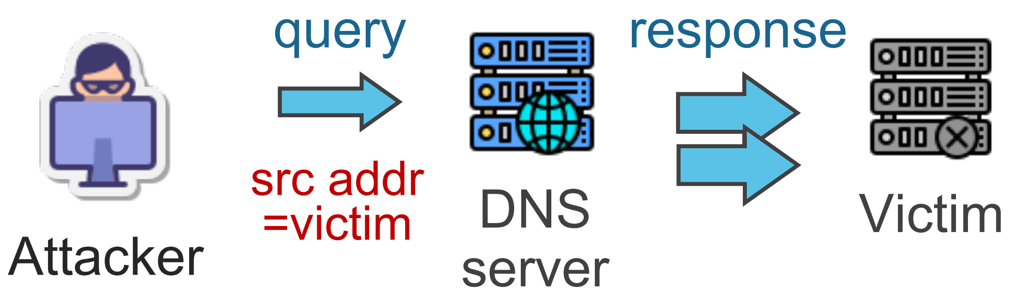
- **Multiple types of attacks have been reported**

| Category | Attack name | Key concept | Amp. factor |
|---|---|---|---|
| **Increasing DNS response size** | Special RRtypes | Exploits large ANY and TXT responses | 200+ |
| | DNSSEC RRs | DNSSEC-signed domains have larger responses | 50+ |
| **Increasing # of DNS responses** | DNS Unchained | Long CNAME chains for resolvers to follow | 8.51 |
| | TsuNAME | Cyclic CNAME/NS dependencies for resolvers to follow | 500 |
| | NXNSAttack | Responses with excessive NSes for resolvers to follow | 3,154 |
| | Routing Loops | Middleboxes in a routing loop intercepting DNS queries | 927,726 * |

* In rare cases only.

# DNS as a classic DoS attack vector

- **Design choices of the DNS protocol**
  - Runs over UDP ➔ *reflected DoS attacks possible*
  - Response larger than query ➔ *traffic amplification*




Reflected DoS attack via DNS

- **Multiple types of attacks have been reported**

| Category | Attack name | Key concept | Amp. factor |
|---|---|---|---|
| **Increasing DNS response size** | Special RRtypes | **Maximizing the amplification potential of *one single DNS server*** | 200+ |
| | DNSSEC RRs | | 50+ |
| **Increasing # of DNS responses** | DNS Unchained | | 8.51 |
| | TsuNAME | | 500 |
| | NXNSAttack | | 3,154 |
| | Routing Loops | | 927,726 * |

* In rare cases only.

# DNS as a classic DoS attack vector

## Even greater DoS potential?

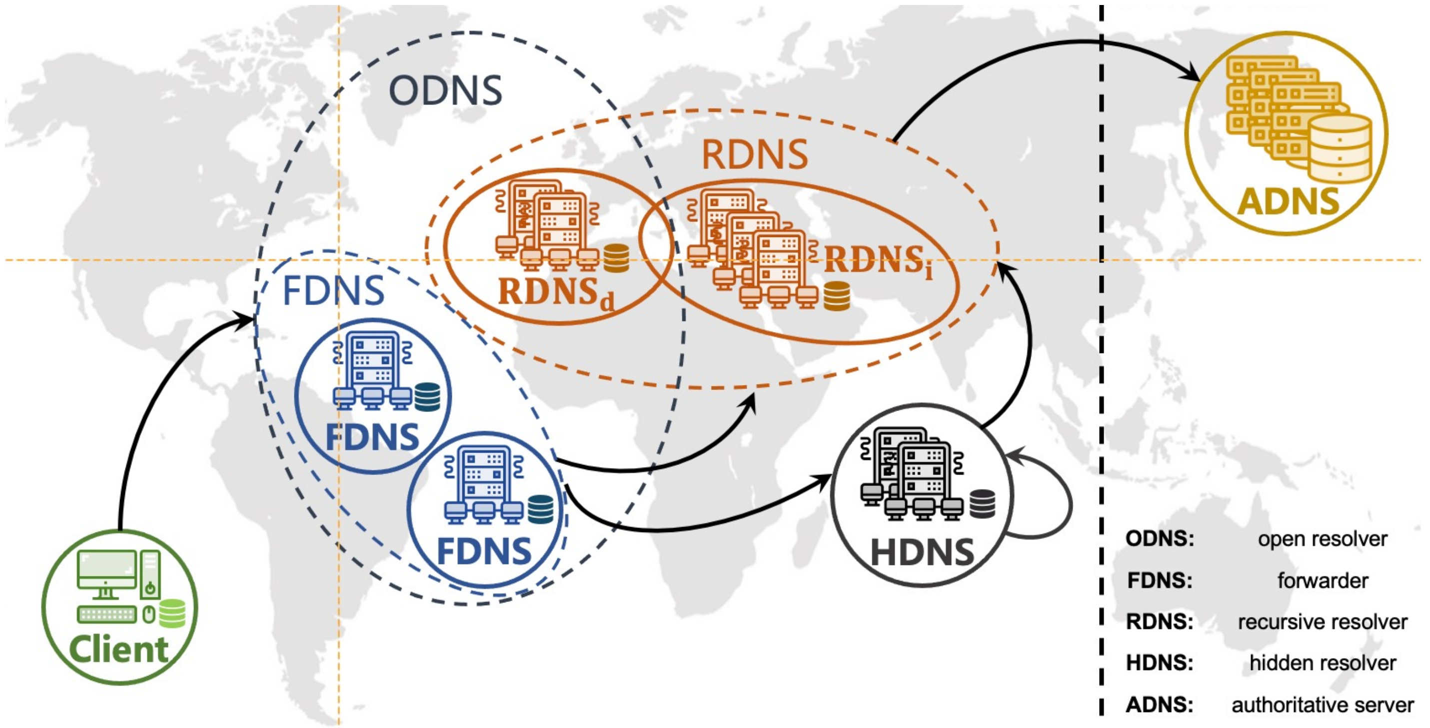
Can we deliberately ***coordinate*** the power of DNS servers to form bigger attacks?

Take a look at how ***complex*** the DNS infrastructure has become.

# DNS as a complex infrastructure

- **Multiple *types* and *layers* of DNS servers**
  - DNS forwarders ➔ pass queries to upstream *(e.g., another forwarder)*
  - Large public DNS services ➔ complexes of load balancers, caches, egress servers, etc.

**The complex DNS infrastructure**

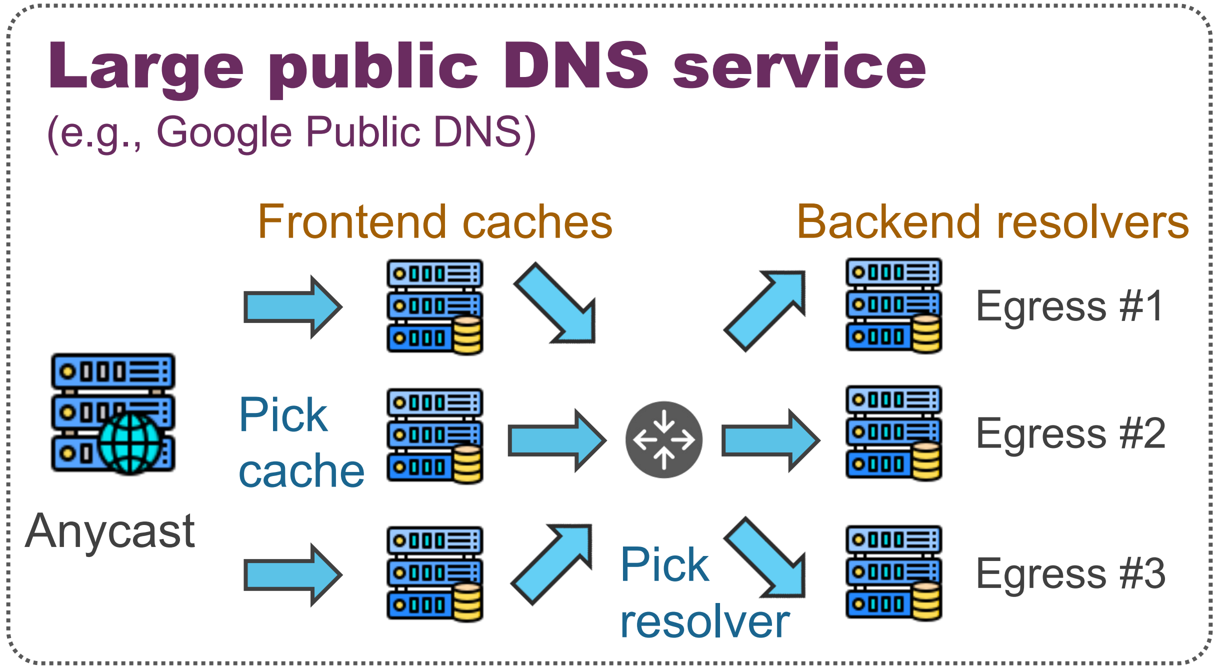



Schomp, et al. On Measuring the Client-side DNS Infrastructure, IMC 2013

**2.27 Million**

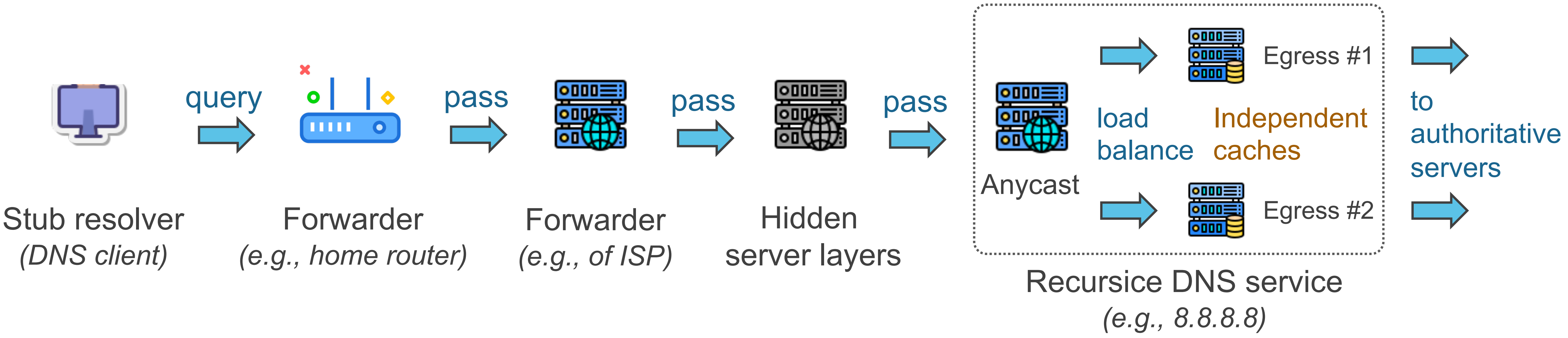
Open DNS servers

* Data from Censys, as of Oct 2023



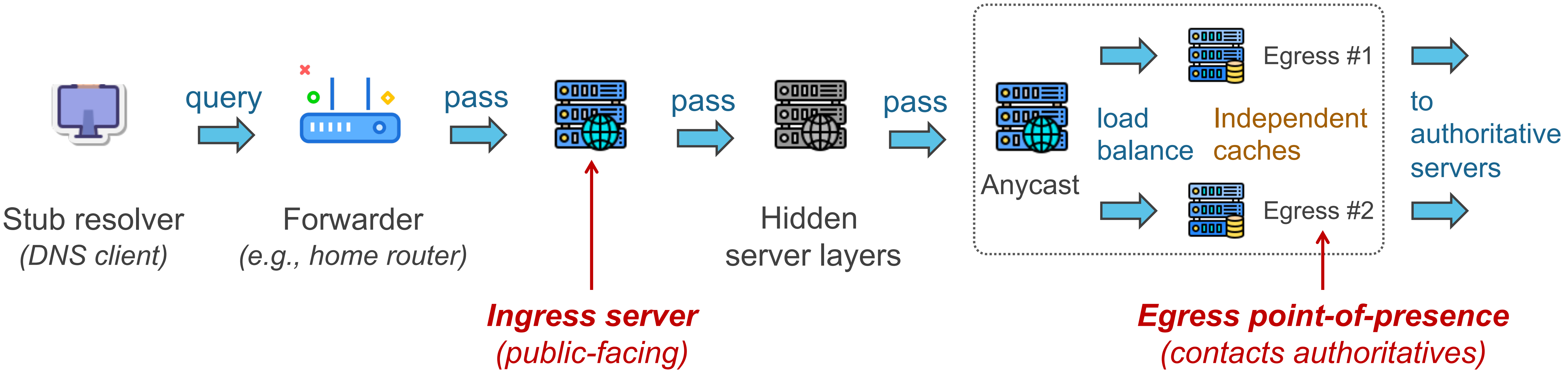
# DNS as a complex infrastructure

- **Multiple *types* and *layers* of DNS servers**
  - DNS forwarders ➔ pass queries to upstream *(e.g., another forwarder)*
  - Large public DNS services ➔ complexes of load balancers, caches, egress servers, etc.
- **A *typical* DNS resolution path now looks like this**



# DNS as a complex infrastructure

- Multiple *types* and *layers* of DNS servers
  - DNS forwarders ➔ pass queries to upstream *(e.g., another forwarder)*
  - Large public DNS services ➔ complexes of load balancers, caches, egress servers, etc.
- A *typical* DNS resolution path now looks like this

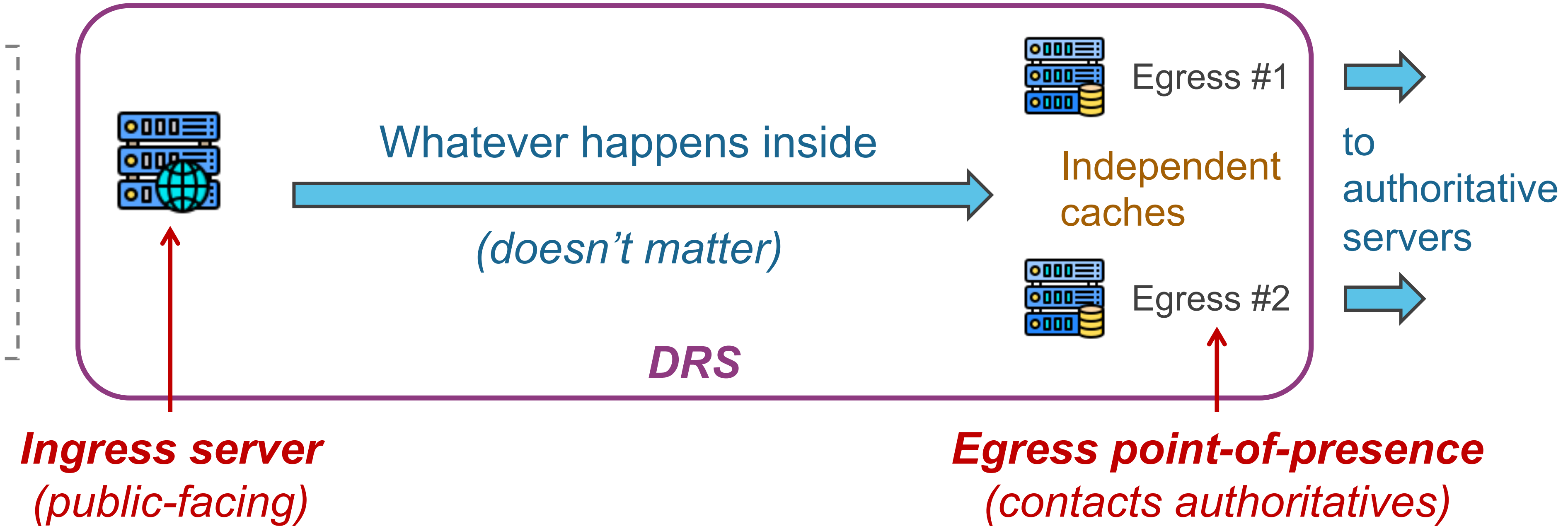


# DNS as a complex infrastructure

- **Multiple *types* and *layers* of DNS servers**
  - DNS forwarders ➔ pass queries to upstream *(e.g., another forwarder)*
  - Large public DNS services ➔ complexes of load balancers, caches, egress servers, etc.
- **A *typical* DNS resolution path now looks like this**

**DNS Resolver System (DRS):**

A public-facing DNS server, together with everything between it and authoritative servers



# DNS as a complex infrastructure

## So I get it, the DNS is complex.

But how is this relevant to
traffic amplifcation?

# Amplification ability: DNS retries

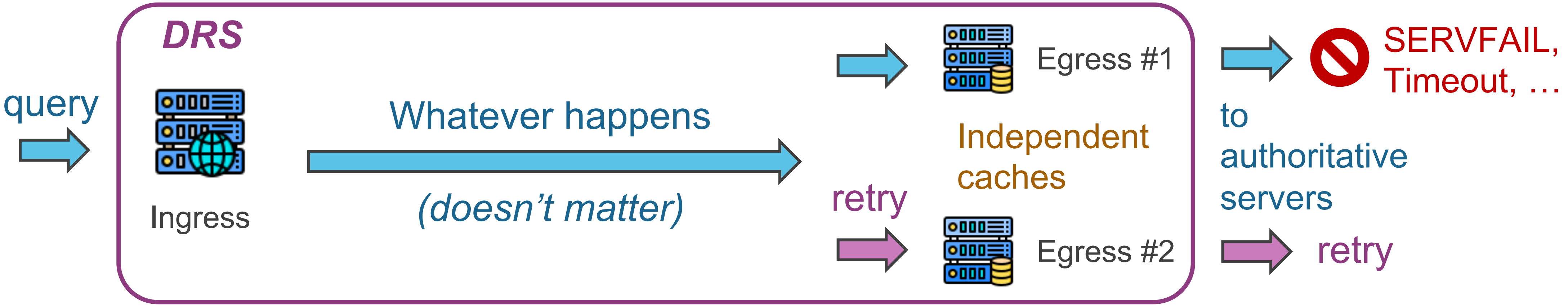
- **DNS is so critical that, it will not take no for an answer**
  - Reasons of DNS failure: *IPv6 incompatibile, timeout, misconfiguration, ...*
- **So upon failure, please *retry* for a few more times**
  - Adopted by mainstream DNS software

| DNS software | # of retries |
|---|---|
| BIND9 | 13 |
| Unbound | 9 |
| Knot | 3 |

# Amplification ability: DNS retries

- **DNS is so critical that, it will not take no for an answer**
  - Reasons of DNS failure: *IPv6 incompatibile, timeout, misconfiguration, ...*
- **So upon failure, please *retry* for a few more times**
  - Adopted by mainstream DNS software
- **For a DRS, retries may exit from *different egresses***
  - Prevents *query aggregation* and *cache hits*

| DNS software | # of retries |
| --- | --- |
| BIND9 | 13 |
| Unbound | 9 |
| Knot | 3 |



# Amplification ability: DNS retries

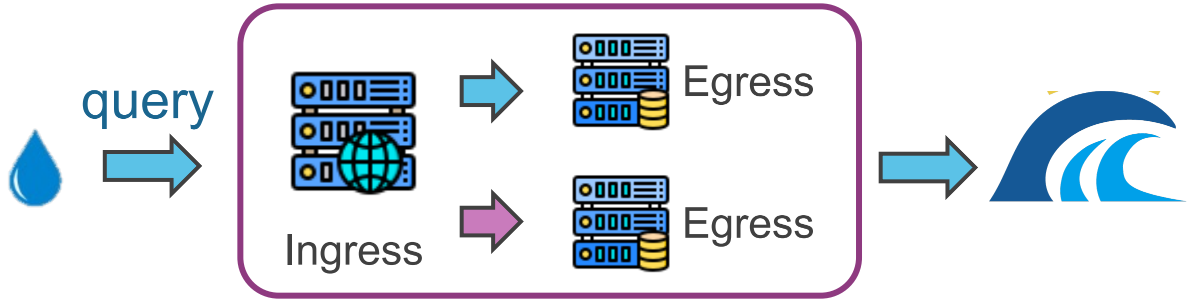
TsuKing

## Wait… You exploit retries?

That's not even enough
to cause ripples!

# Attack variant I: DNSRetry

- **There are bogus DRS implementations that retry aggressively**
  - They themselves already are powerful amplifiers
  - Max # of retries by one DRS: ***117,541***

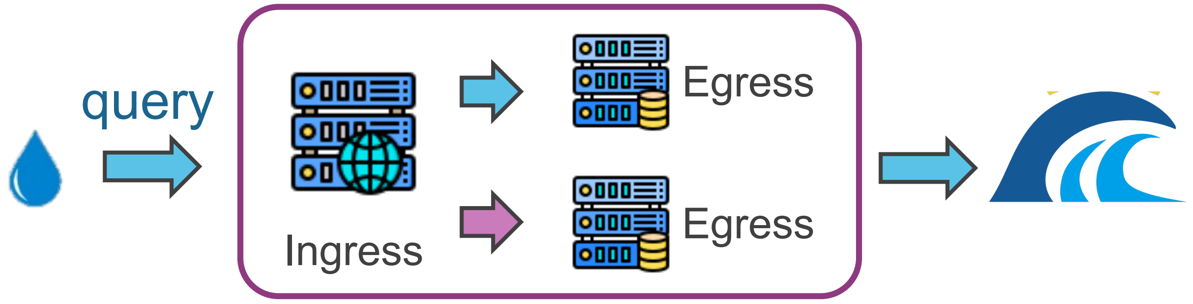



| # of retries | # of open DRSes | % of tested |
|---|---|---|
| > 2 | 925,500 | 69.8% |
| > 10 | 407,581 | 30.7% |
| > 100 | 31,660 | 2.4% |
| **> 1,000** | **529** | **0.04%** |

- **How to divert all retries to one victim?**
  - Using *referrals*; explained soon.

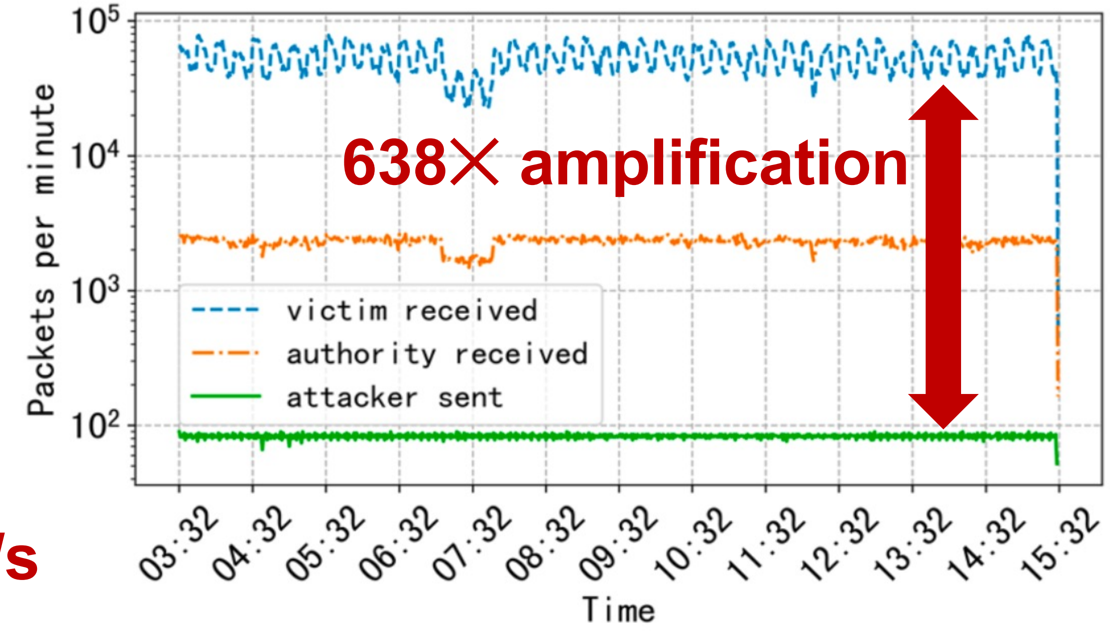
# Attack variant I: DNSRetry - Evaluation

- **There are bogus DRS implementations that retry aggressively**
  - They themselves already are powerful amplifiers
  - Max # of retries by one DRS: ***117,541***




| # of retries | # of open DRSes | % of tested |
|---|---|---|
| > 2 | 925,500 | 69.8% |
| > 10 | 407,581 | 30.7% |
| > 100 | 31,660 | 2.4% |
| **> 1,000** | **529** | **0.04%** |

- **How to divert all retries to one victim?**
  - Using *referrals*; explained soon.
- **Evaluation in controlled environment**
  - Select 10 DRSes that retry aggresively
  - Attacker sends 1.3 pkt/s ➔ **Victim receives 882 pkt/s**



# Attack variant I: DNSRetry - Evaluation

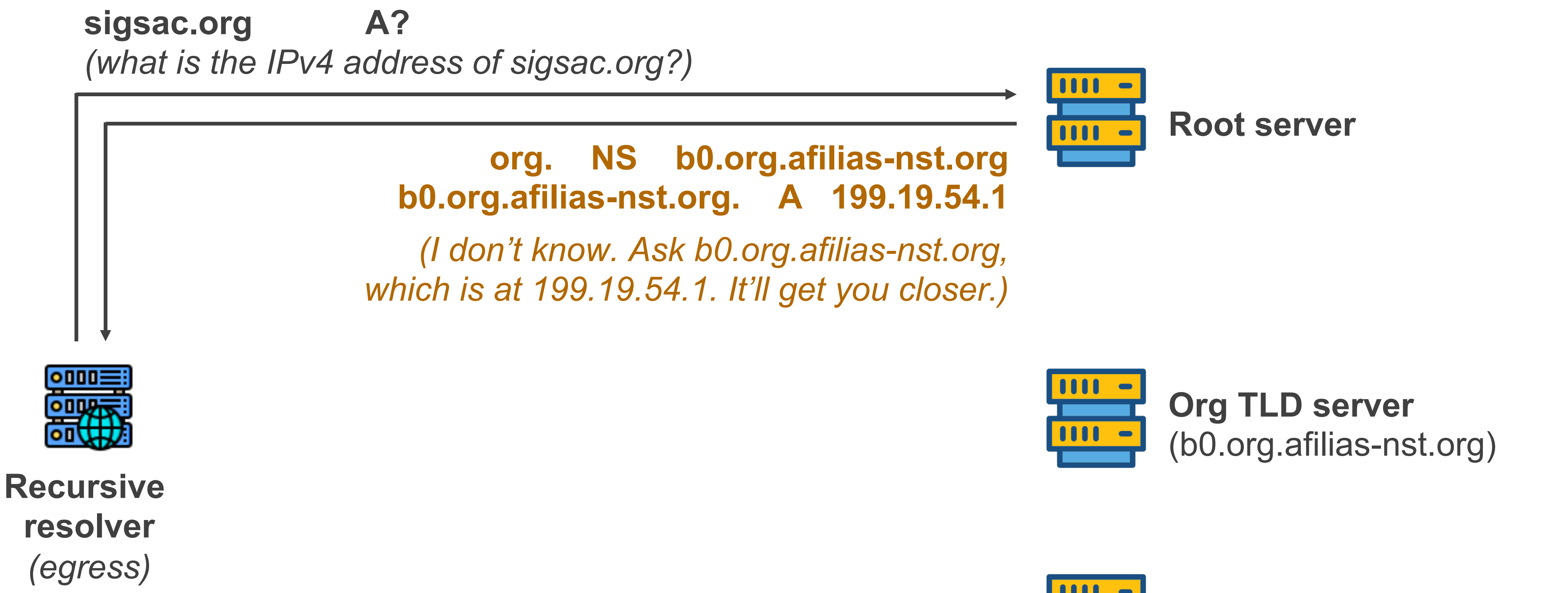
TsuKing

## Alright, but lots of them are not aggressive at all…

Let’s ***chain*** these ripples into bigger waves!

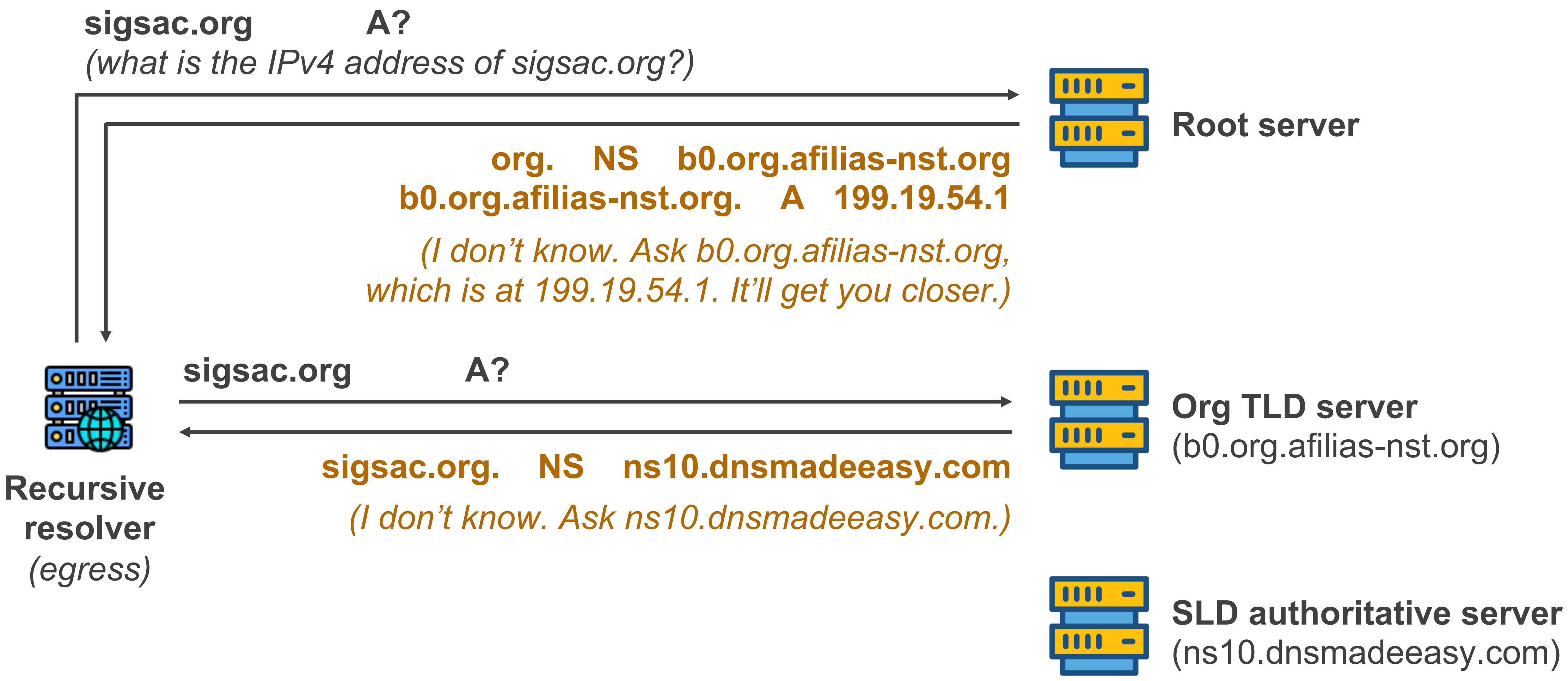
# Coordination ability: Referrals

- **Recursive DNS resolution guided by *referrals***
  - Referrals *tell recursive resolvers who to ask next*



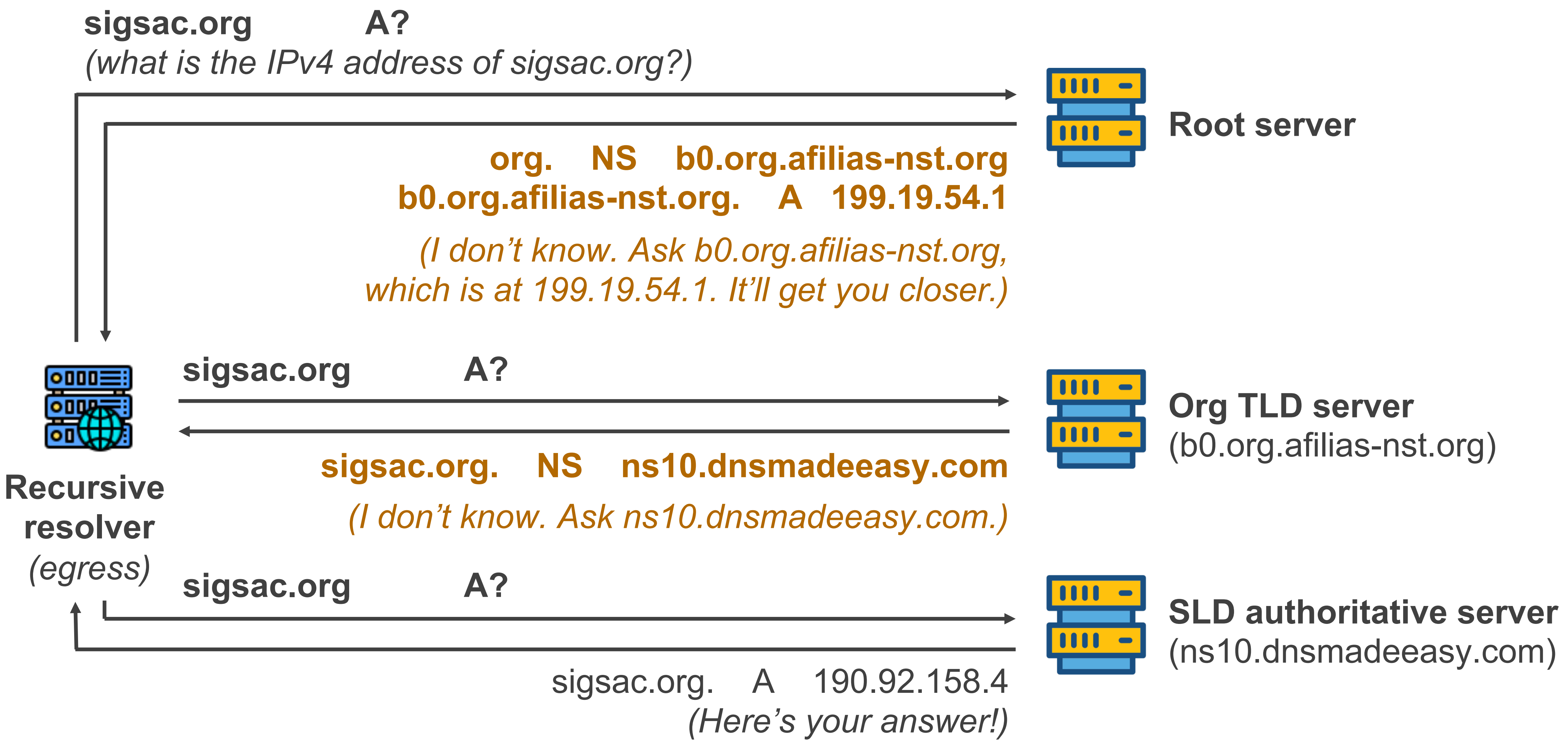
# Coordination ability: Referrals

- **Recursive DNS resolution guided by *referrals***
  - Referrals *tell recursive resolvers who to ask next*



# Coordination ability: Referrals

- Recursive DNS resolution guided by ***referrals***
  - Referrals *tell recursive resolvers who to ask next*

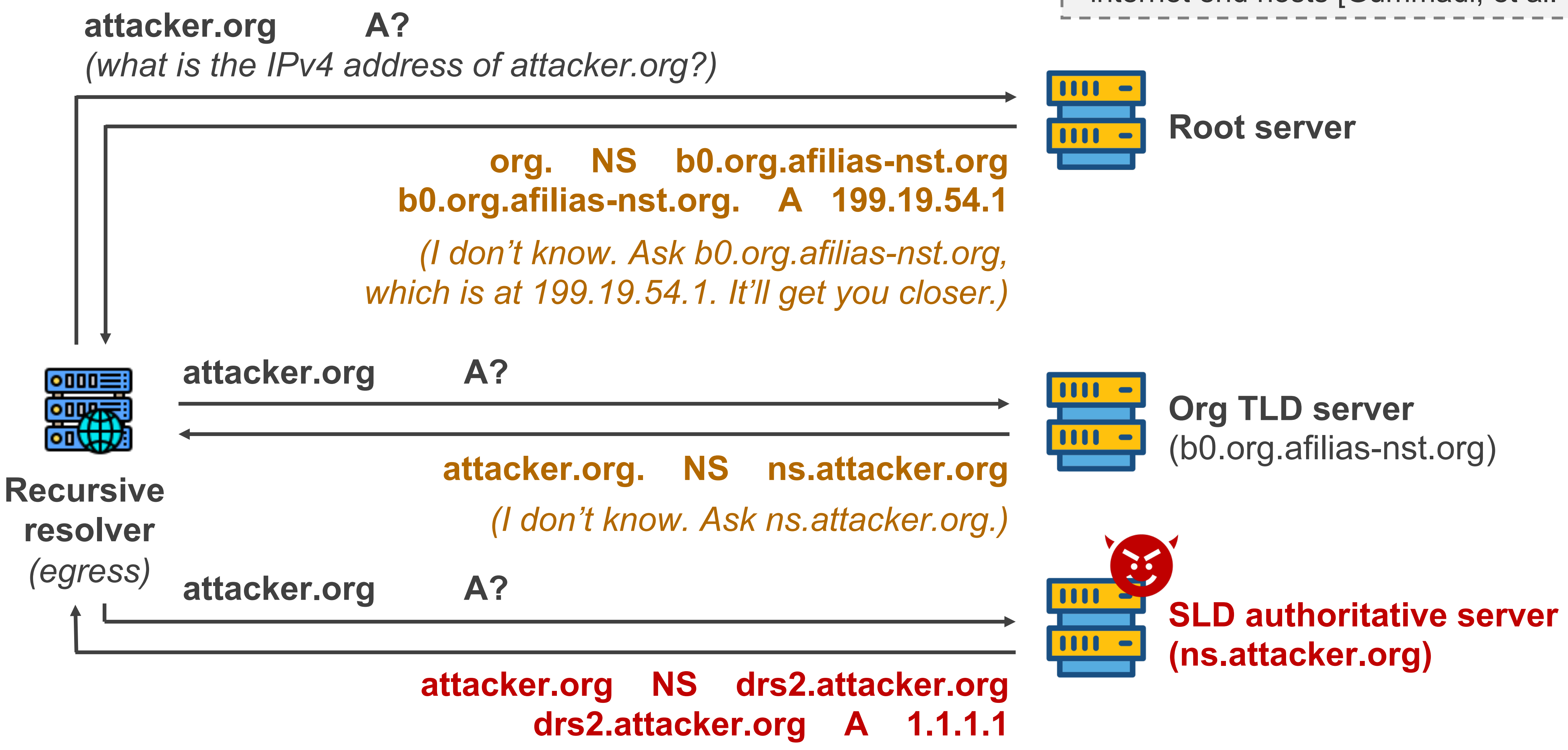


# Coordination ability: Referrals

- ❖ **Recursive DNS resolution guided by *referrals***
  - ❖ Use ***evil referrals*** to divert queries arbitrarilly

**Inspired by:**

**King**: estimating latency between arbitrary internet end hosts [Gummadi, et al. CCR ’02]



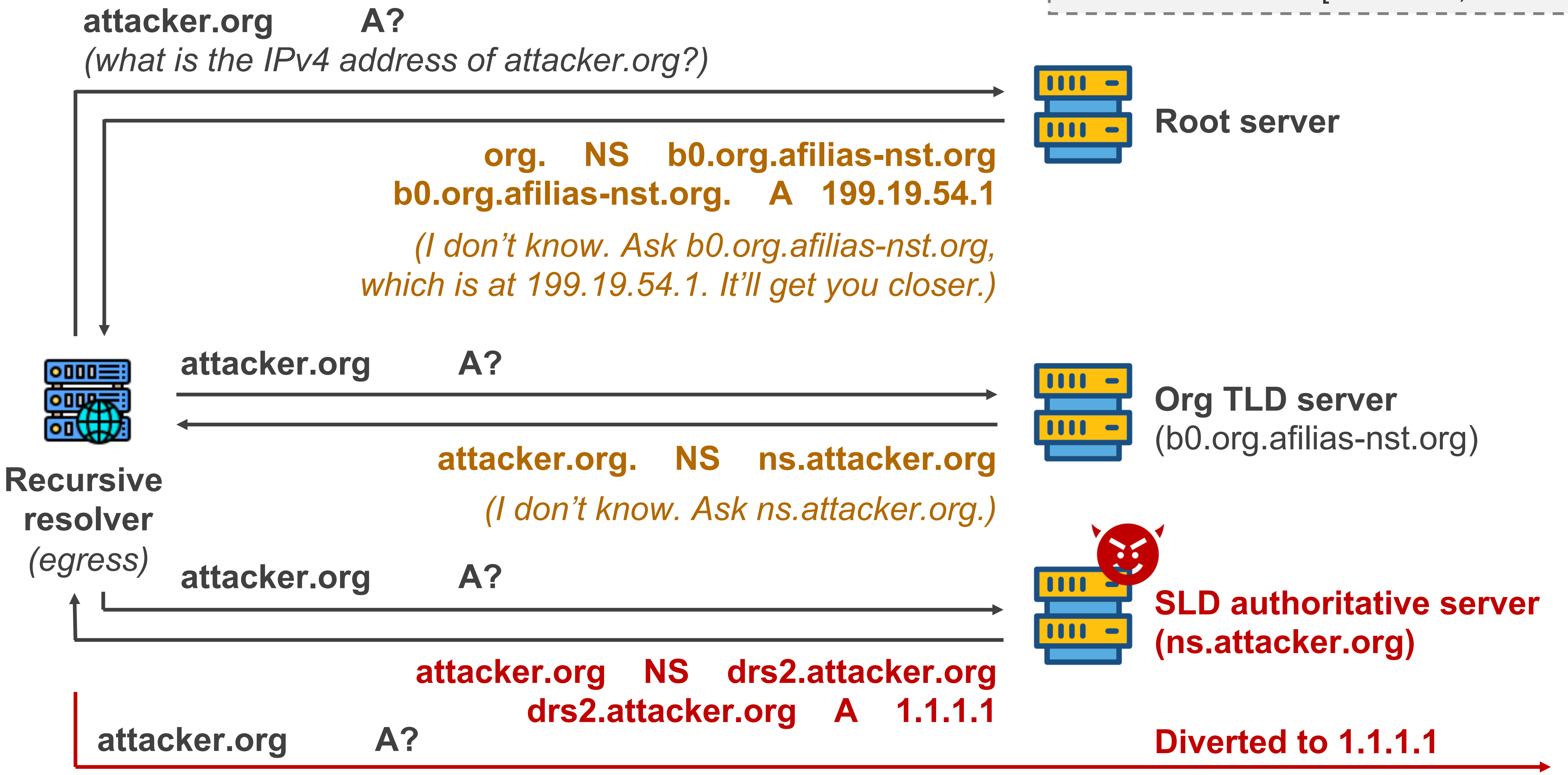
# Coordination ability: Referrals

TsuKing

- Recursive DNS resolution guided by ***referrals***
  - Use ***evil referrals*** to divert queries arbitrarilly

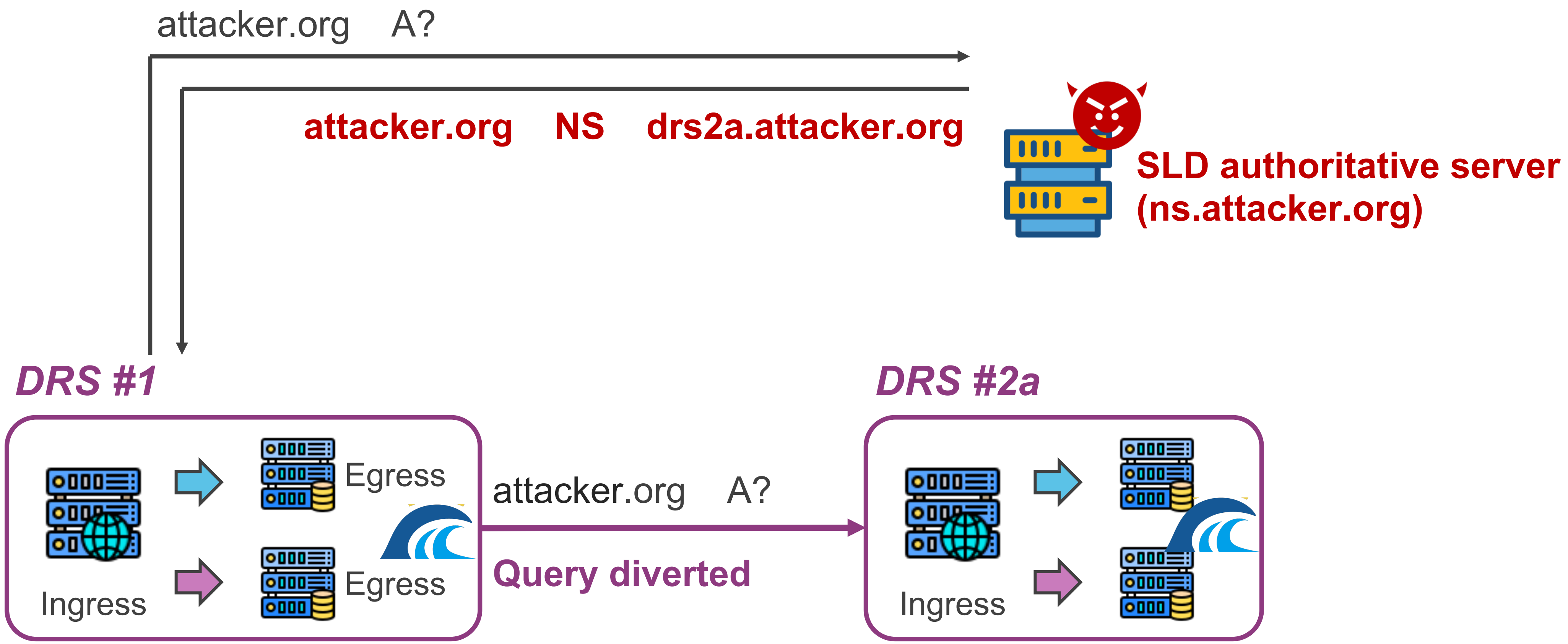
**Inspired by:**

**King**: estimating latency between arbitrary internet end hosts [Gummadi, et al. CCR ’02]



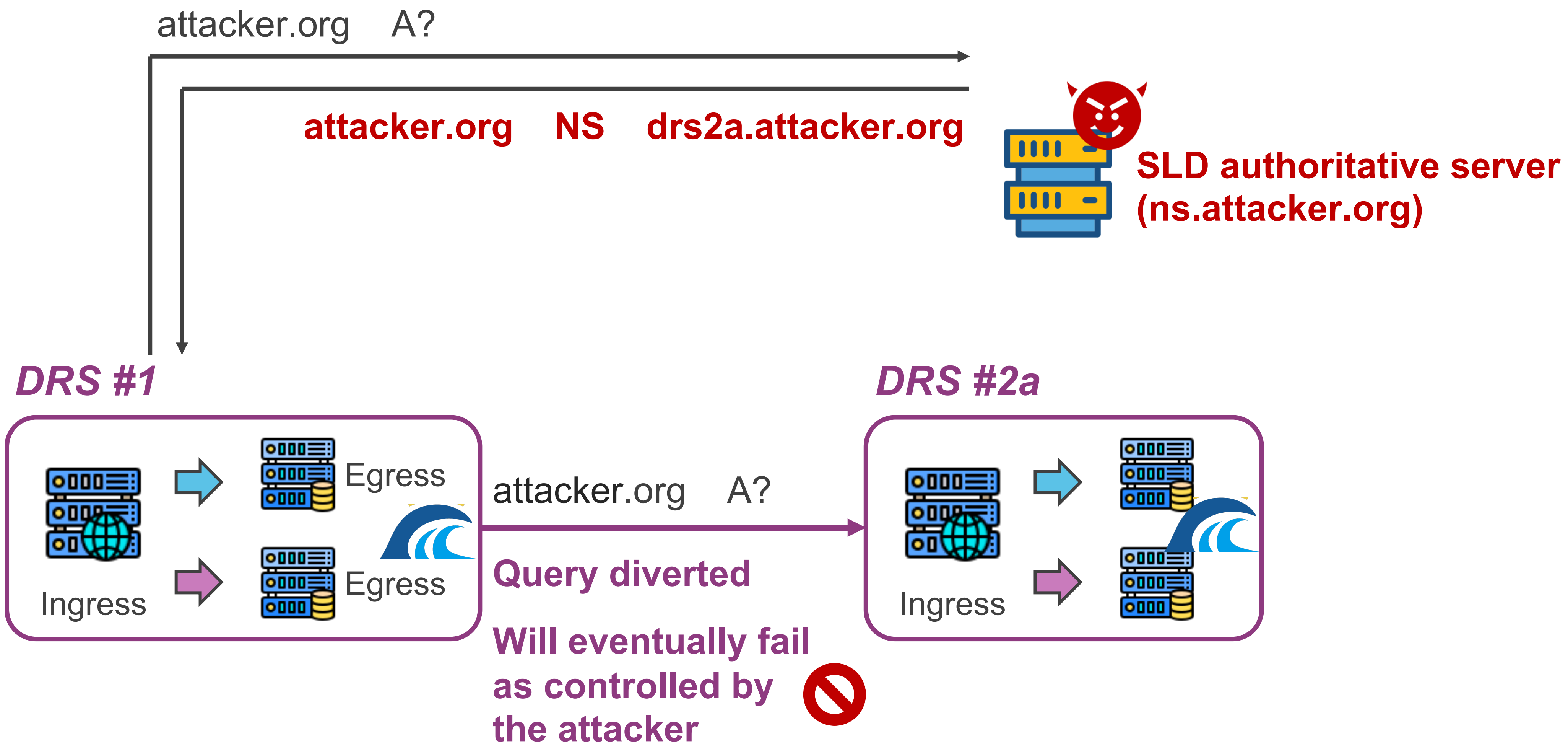
# Attack variant II: DNSChain

❖ Recursive DNS resolution guided by *evil referrals*



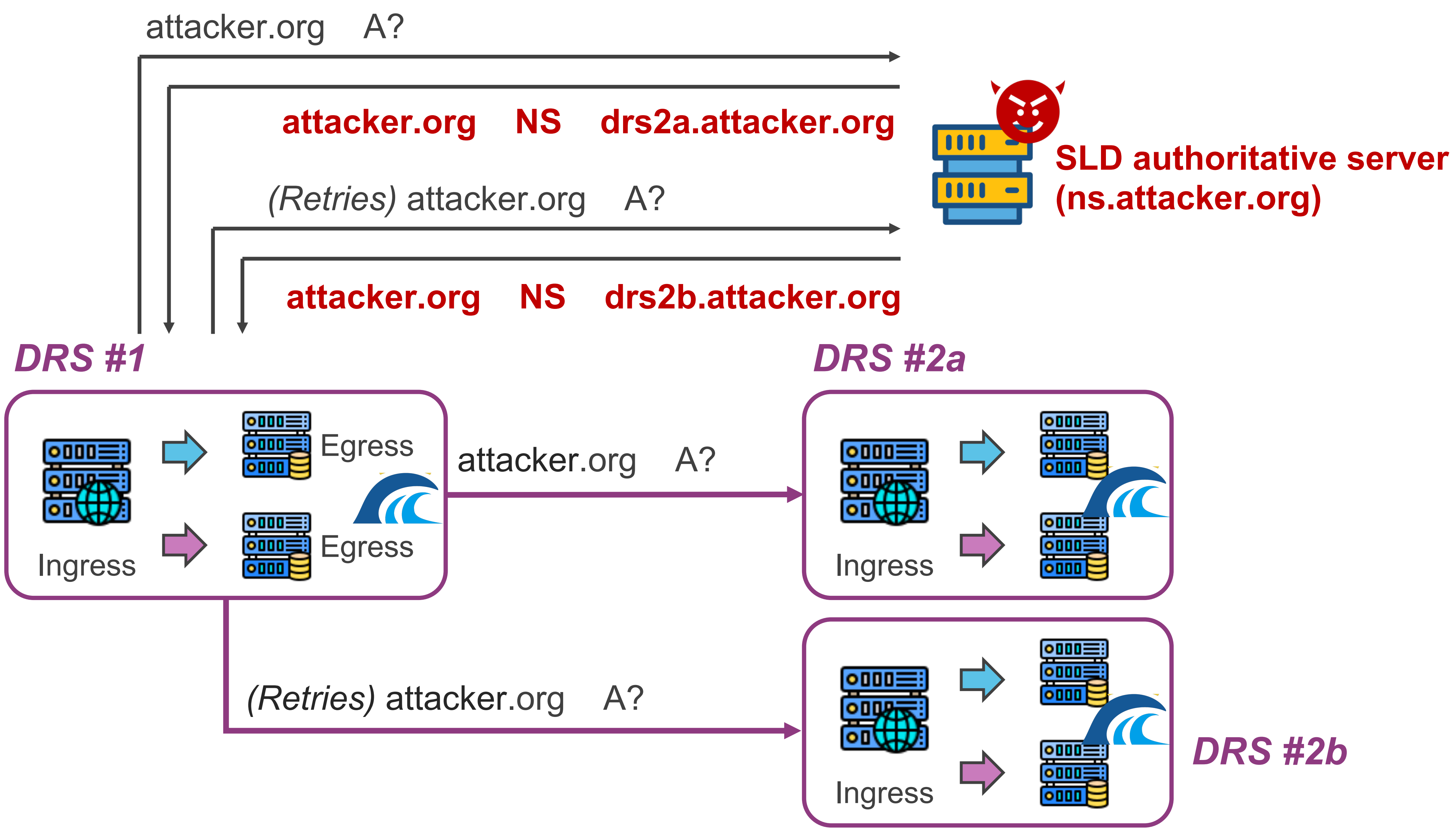
# Attack variant II: DNSChain

❖ Recursive DNS resolution guided by ***evil referrals***



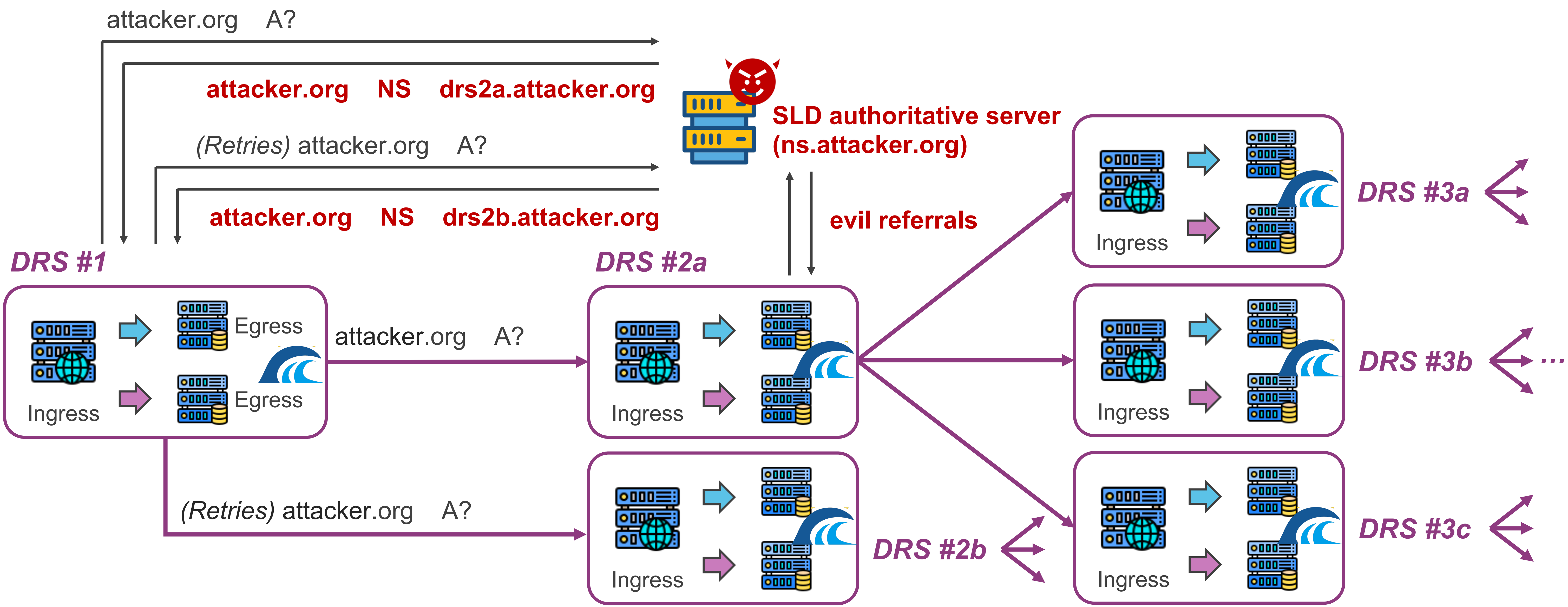
# Attack variant II: DNSChain

❖ Recursive DNS resolution guided by ***evil referrals***



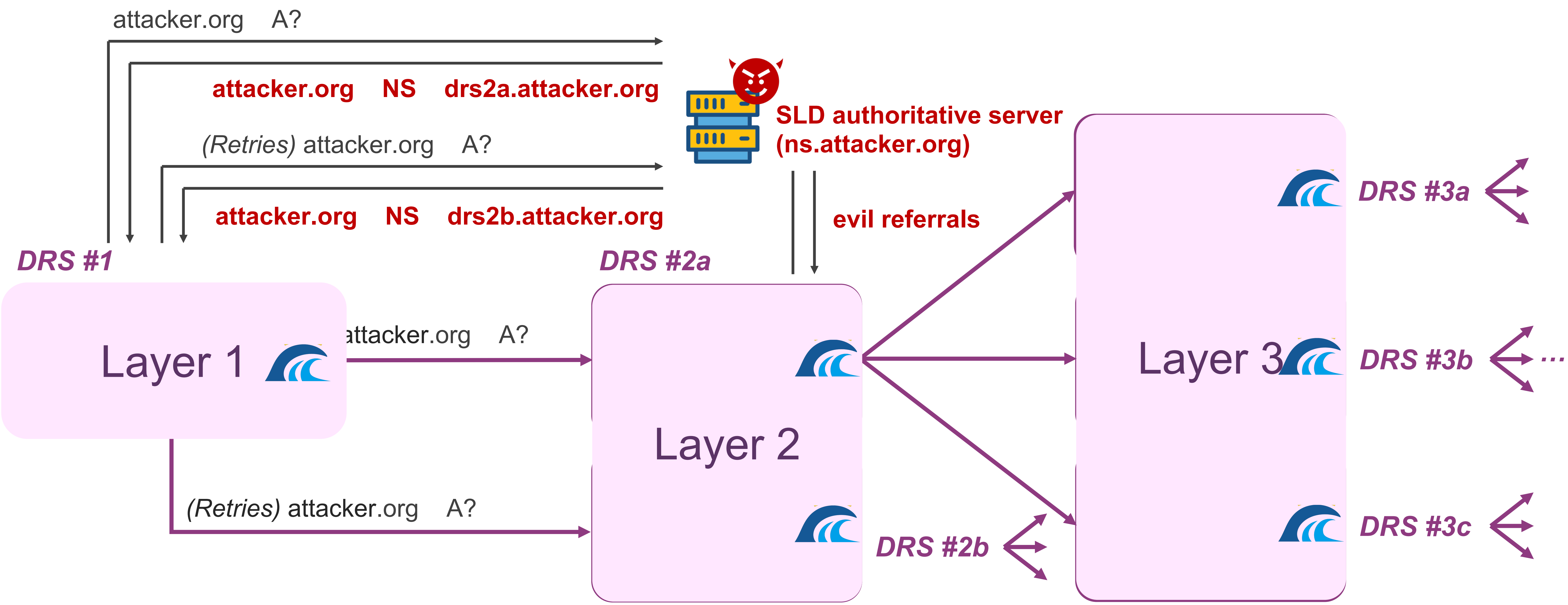
# Attack variant II: DNSChain

❖ Recursive DNS resolution guided by ***evil referrals***



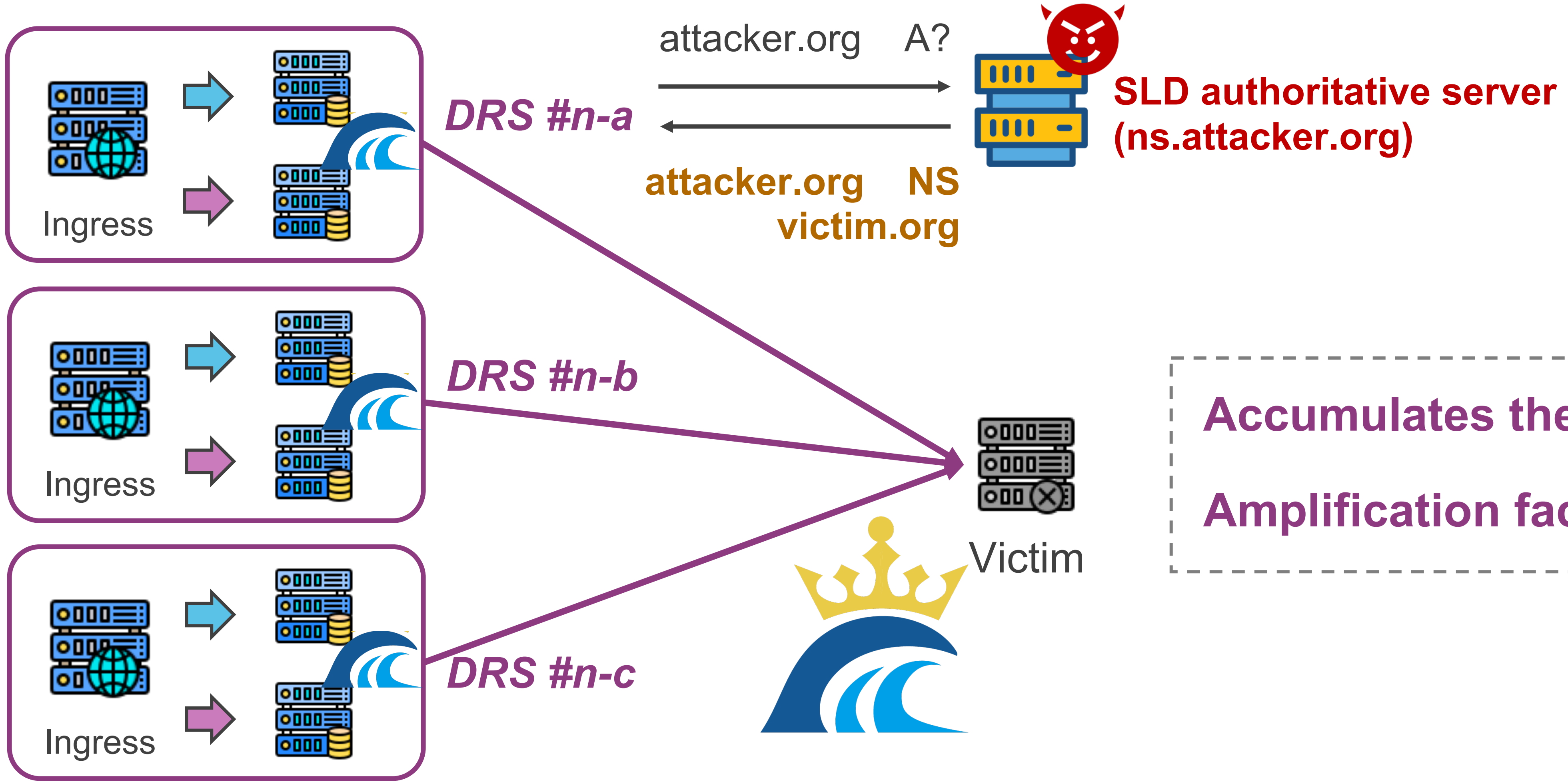
# Attack variant II: DNSChain

❖ **Recursive DNS resolution guided by *evil referrals***



# Attack variant II: DNSChain

- **Recursive DNS resolution guided by *evil referrals***
  - ***Final referral:*** points to victim




**Accumulates the power of layers of DRSes**

**Amplification factor *multiplies***

# Attack variant II: DNSChain - Evaluation

- **Evaluation in controlled environment**
  - We select from exploitable DRSes and coordinate them into ***layers***

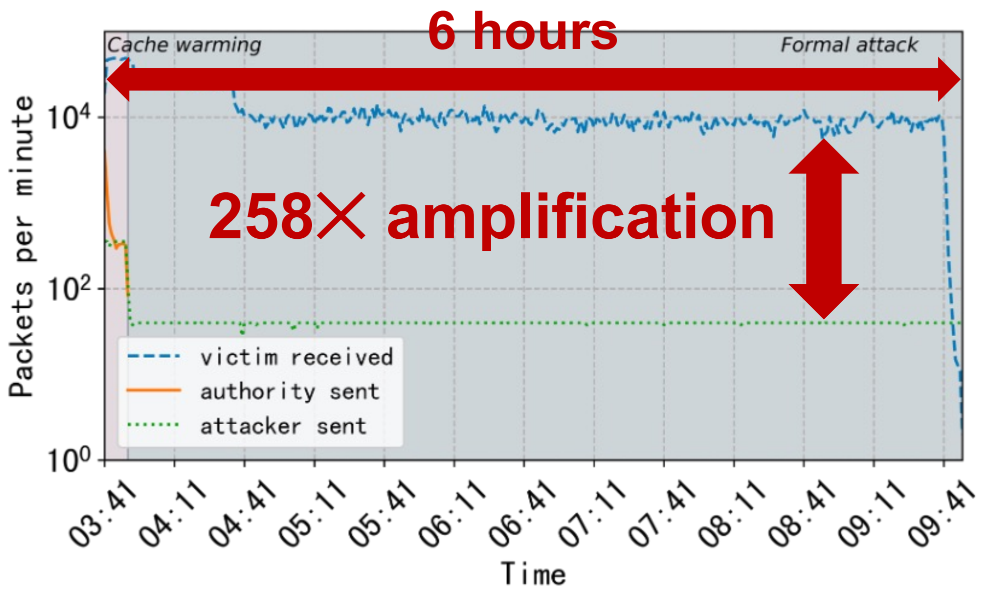
| Setting | # of DRSes coordinated in each layer | | | | | | | Amp. factor |
|---|---|---|---|---|---|---|---|---|
| | Layer 1 | Layer 2 | Layer 3 | Layer 4 | Layer 5 | Layer 6 | Layer 7 | |
| **# 1** | 1 | 4 | 8 | - | - | - | - | 288 |
| **# 2** | 1 | 4 | 8 | 16 | 32 | - | - | 591 |
| **# 3** | 1 | 4 | 8 | 16 | 32 | 64 | 128 | **3,702** |

# Attack variant II: DNSChain - Evaluation

- **Evaluation in controlled environment**
  - We select from exploitable DRSes and coordinate them into ***layers***

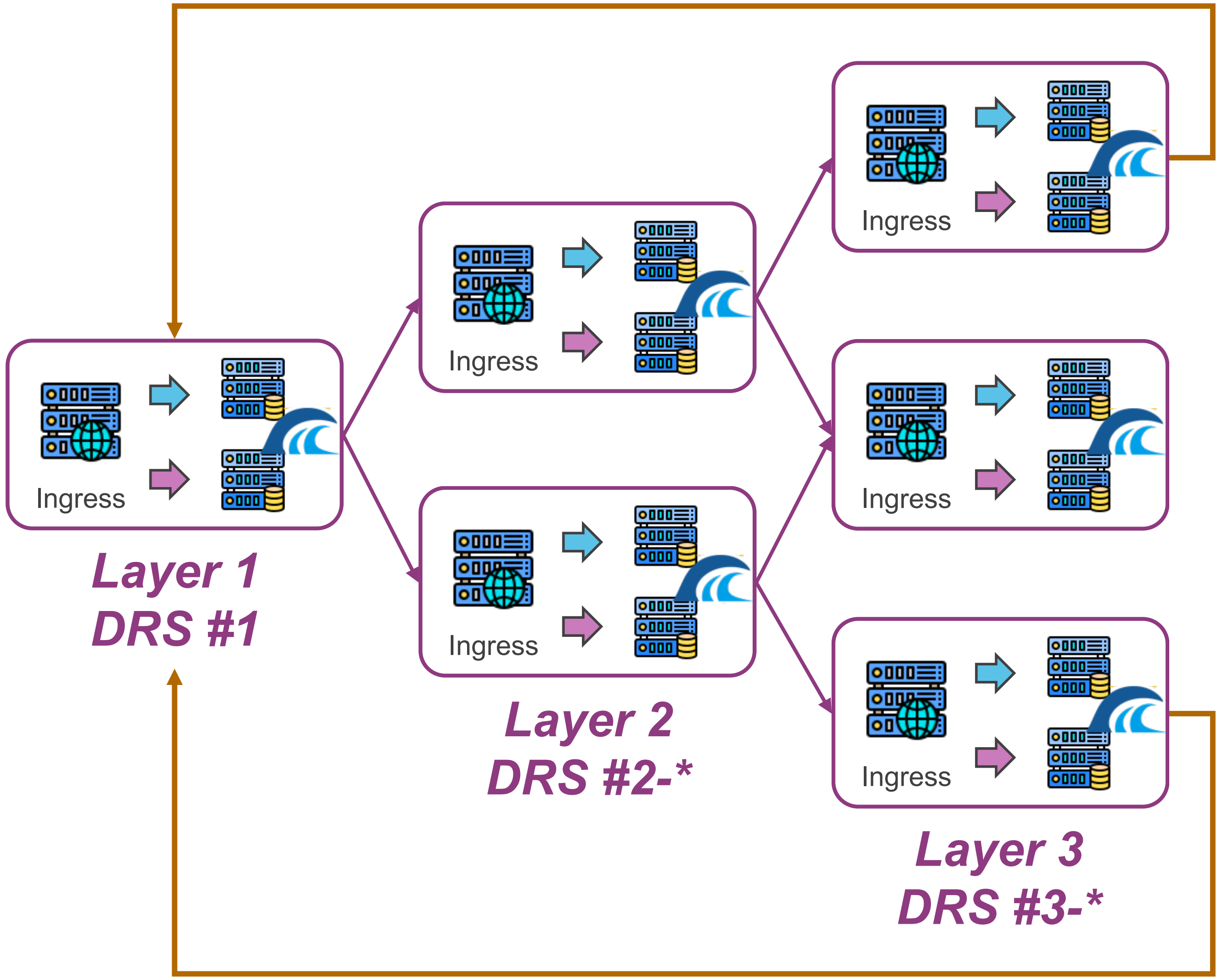
| Setting | # of DRSes coordinated in each layer | | | | | | | Amp. factor |
|---|---|---|---|---|---|---|---|---|
| | Layer 1 | Layer 2 | Layer 3 | Layer 4 | Layer 5 | Layer 6 | Layer 7 | |
| **# 1** | 1 | 4 | 8 | - | - | - | - | 288 |
| **# 2** | 1 | 4 | 8 | 16 | 32 | - | - | 591 |
| **# 3** | 1 | 4 | 8 | 16 | 32 | 64 | 128 | **3,702** |

- **Can the attack last?**
  - Setting #2 (5 layers); attacker send at 0.8 pkt/s
  - ***Amplification effect persists in 6 hours***



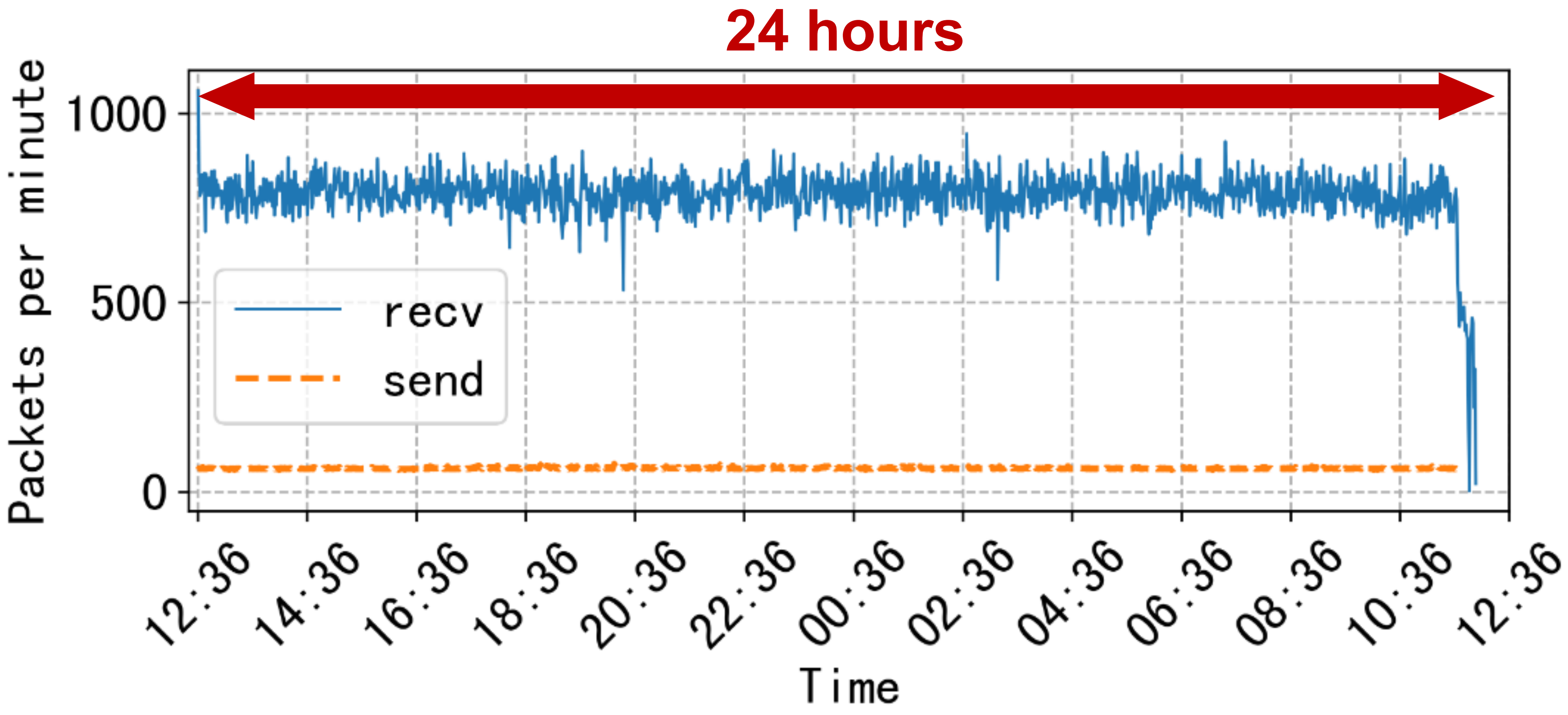
# Attack variant III: DNSLoop

- **Modified from DNSChain, creating a *loop* of retry queries**
  - ***Final referral:*** points back to DRS #1
- **The victim and goal change now**
  - ***ALL DRSes in the loop*** become victims
  - Goal is to exhause their resources
  - *Increasing amplification factor is a non-goal*
- **Attackers may also**
  - Inject new rounds of retries to the loop
  - Simply by querying DRS #1



# Attack variant III: DNSLoop - Evaluation

- **Evaluation in controlled environment - can the loop last?**
  - Coordinates 7 layers of DRSes
  - Build RouterOS host as ingress *(rate limit at 1 pkt/s, due to ethical considerations)*
  - ***Attacker sends 1 query only, loop lasts until deliberate stop***



# Attack variants II & III

## Seems overwhelming, but can many DRSes be used?

What are the conditions of successful attacks?

# Conditions of successful attacks

TsuKing

- **DRS *not honoring cleared RD bit* in DNS header**
  - RD (recursion desired) =0: *do not perform recursion, find answers locally in cache*
  - Usually *cleared by egress*, as authoritative servers cannot perform recursion
  - DRS honors RD ➔ *chain cannot continue*
  - ***27.2% of tested DRSes do not honor***

| | | | | | | |
|---|---|---|---|---|---|---|
| Transaction ID | QR | Opcode | RD | Flags | Z | RCODE |
| QDCOUNT | | | ANCOUNT | | | |
| NSCOUNT | | | ARCOUNT | | | |

# Conditions of successful attacks

TsuKing

- **DRS *not honoring cleared RD bit* in DNS header**
  - RD (recursion desired) =0: *do not perform recursion, find answers locally in cache*
  - Usually *cleared by egress*, as authoritative servers cannot perform recursion
  - DRS honors RD ➔ *chain cannot continue*
  - ***27.2% of tested DRSes do not honor***

| | | | | | | |
|---|---|---|---|---|---|---|
| Transaction ID | QR | Opcode | RD | Flags | Z | RCODE |
| QDCOUNT | | | ANCOUNT | | | |
| NSCOUNT | | | ARCOUNT | | | |

- **DRS not deployed with negative caching [RFC 2308]**
  - Negative caching records DNS failures ➔ *effectively eliminates retries*
  - ***43% of tested DRSes do not deploy***

# Conditions of successful attacks

- **DRS *not honoring cleared RD bit* in DNS header**
  - RD (recursion desired) =0: *do not perform recursion, find answers locally in cache*
  - Usually *cleared by egress*, as authoritative servers cannot perform recursion
  - DRS honors RD ➔ *chain cannot continue*
  - ***27.2% of tested DRSes do not honor***

| Transaction ID | QR | Opcode | RD | Flags | Z | RCODE |
|---|---|---|---|---|---|---|
| QDCOUNT | ANCOUNT | | | | | |
| NSCOUNT | ARCOUNT | | | | | |

- **DRS not deployed with negative caching [RFC 2308]**
  - Negative caching records DNS failures ➔ *effectively eliminates retries*
  - ***43% of tested DRSes do not deploy***

- **DRS has multiple egresses: *the more, the better***
  - ***52% of tested DRSes have over 10 egresses***

# Mitigation

## What can we do to prevent this?

Correct bogus implementations such that attack conditions cannot be fulfilled.

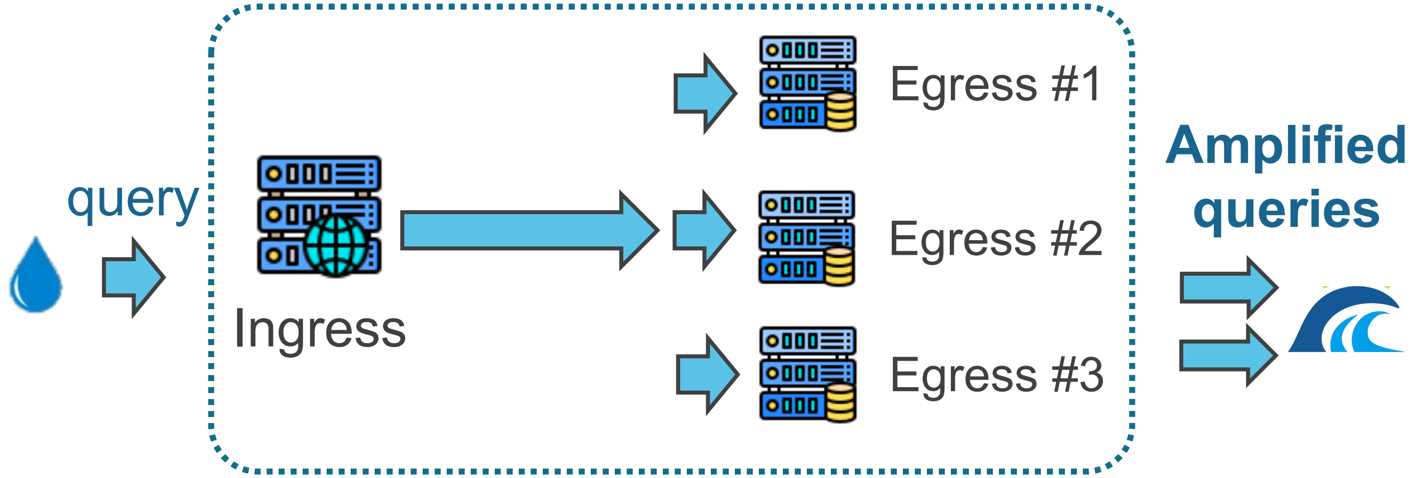
# Mitigation

## Tsu-King

### Tsunami

*(Traffic amplification ability)*

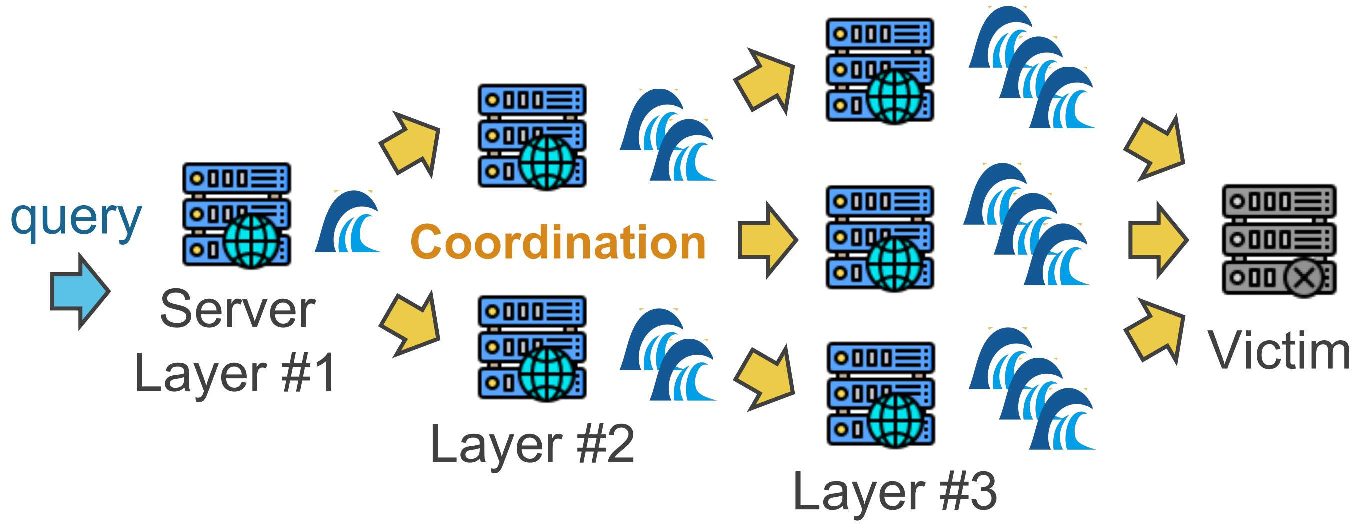
- Cause 1: complex DNS infrastructure
- Cause 2: aggressive retries exhibited by bogus implementations




### King

*(Server coordination ability)*

- Cause 3: not following DNS specifications (in this case, the *RD* flag)



# Mitigation

- **Avoid aggressive retries**
  - A modest number of retries should suffice, as adopted by mainstream software
- **Follow DNS specifications**
  - Honor the DNS flags: if RD tells not to perform recursion, just don't
- **Deploy additional mechanisms that add protection**
  - Negative caching: good to reduce retries
  - Egress and cache management: reduce independence between egress servers

# Feedback from vendors

- **DNS software & public DNS: *not honoring RD flag***
  - Confirmed and fixed: *RouterOS, Unbound; 114DNS, AliDNS, DNSPod*
  - Proposed plans but not accepted as security issue: *PowerDNS*

> ✓ **- Fix not following cleared RD flags potentially enables amplification**
> DDoS attacks, reported by Xiang Li and Wei Xu from NISL Lab,
> Tsinghua University. The fix stops query loops, by refusing to send
> RD=0 queries to a forwarder, they still get answered from cache.

**Unbound fix message**

- **3 assigned CVE enrties**

**CVE-2023-24711** **CVE-2023-24712** **CVE-2023-28455**

# Questions?

**Paper website:** **https://tsuking.net**

**Lead authors:**


- ❖ Wei Xu (xu-w21@mails.tsinghua.edu.cn)
- ❖ Xiang Li (x-l19@mails.tsinghua.edu.cn)


**Presented by:**


- ❖ Chaoyi Lu (luchaoyi@tsinghua.edu.cn, https://chaoyi.lu)