Measuring Query Latency of Top Level DNS Servers

Jinjin Liang[1], Jian Jiang[1], Haixin Duan[1], Kang Li[2] and Jianping Wu[1]

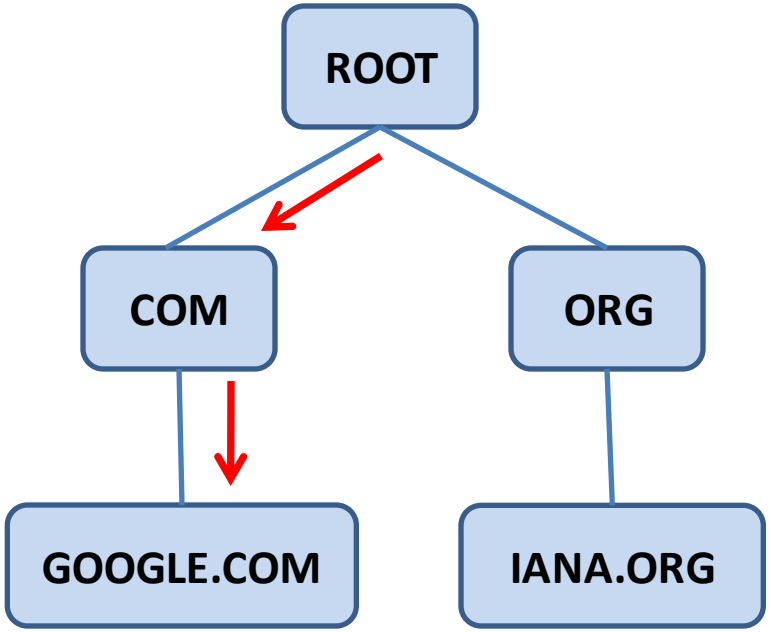
Tsinghua University, China[1]
University of Georgia[2]

PAM 2013

DNS Overview

- Domain Name System
  - Translate domain names to IP addresses
  - Initial step for most Internet applications

- Top Level Zones
  - Start points of resolutions
  - Even with local cache



Replication: State of the art

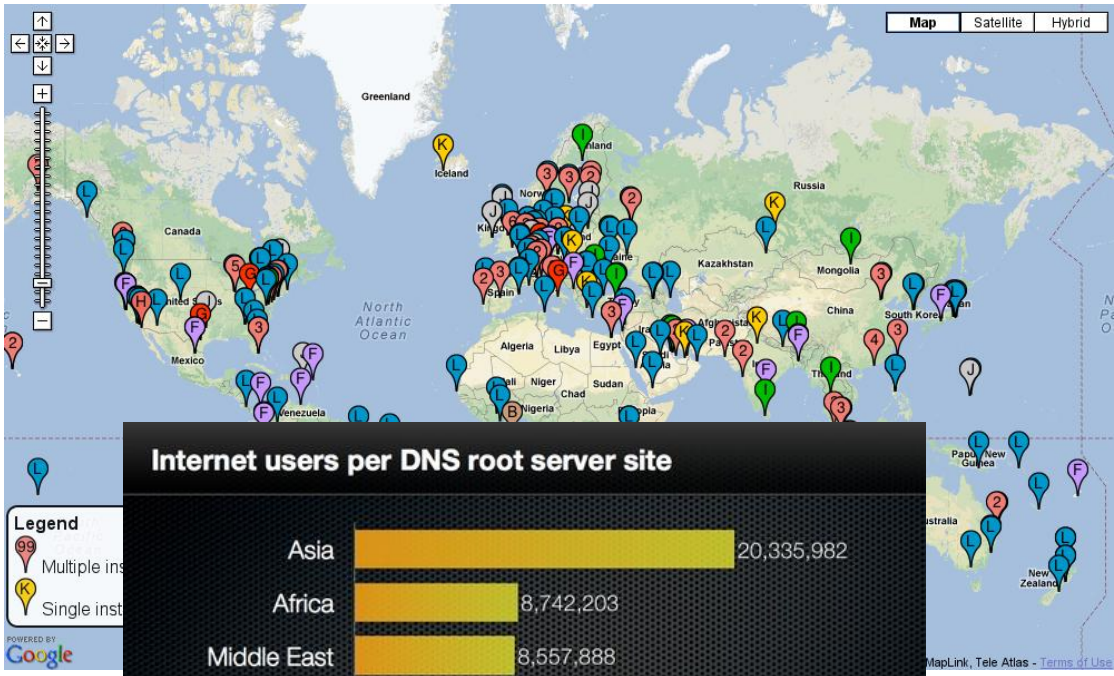
- Root Zone
  - Zone Replications
    - 13 Roots (A~M)
    - Uneven QoS

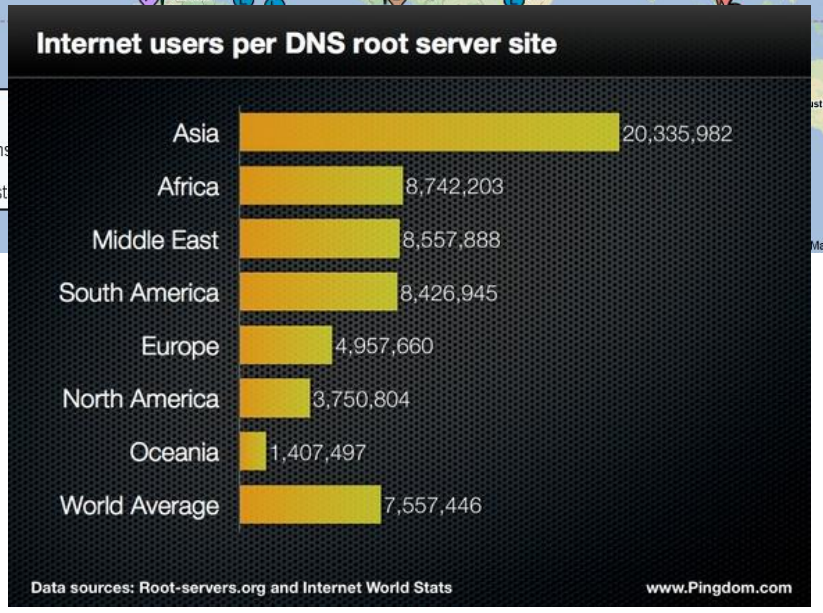
Replication: State of the art

- Root Zone
  - Zone Replications
    - 13 Roots (A~M)
    - Uneven QoS
  - Anycast
    - 319 instances
    - All over the world

Replication: State of the art

- Root Zone
  - Zone Replications
    - 13 Roots (A~M)
    - Uneven QoS
  - Anycast
    - 319 instances
    - All over the world

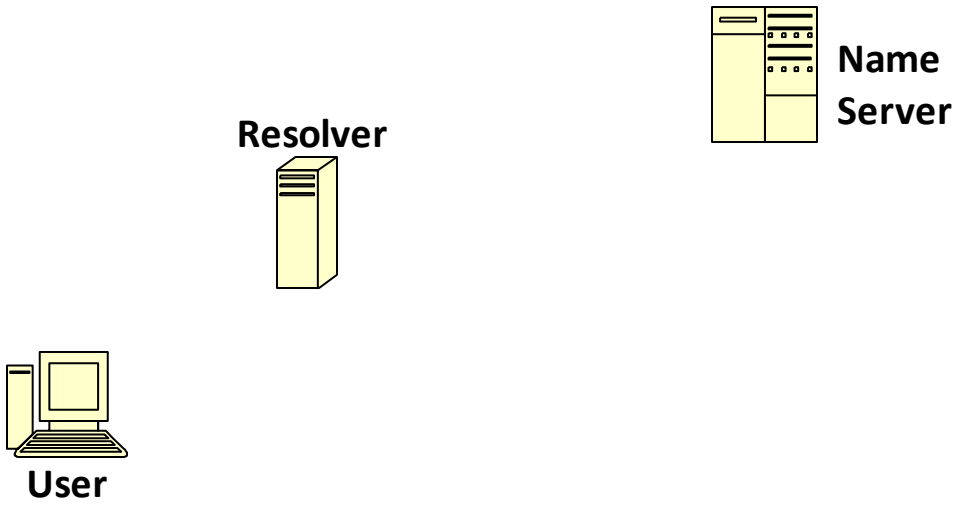




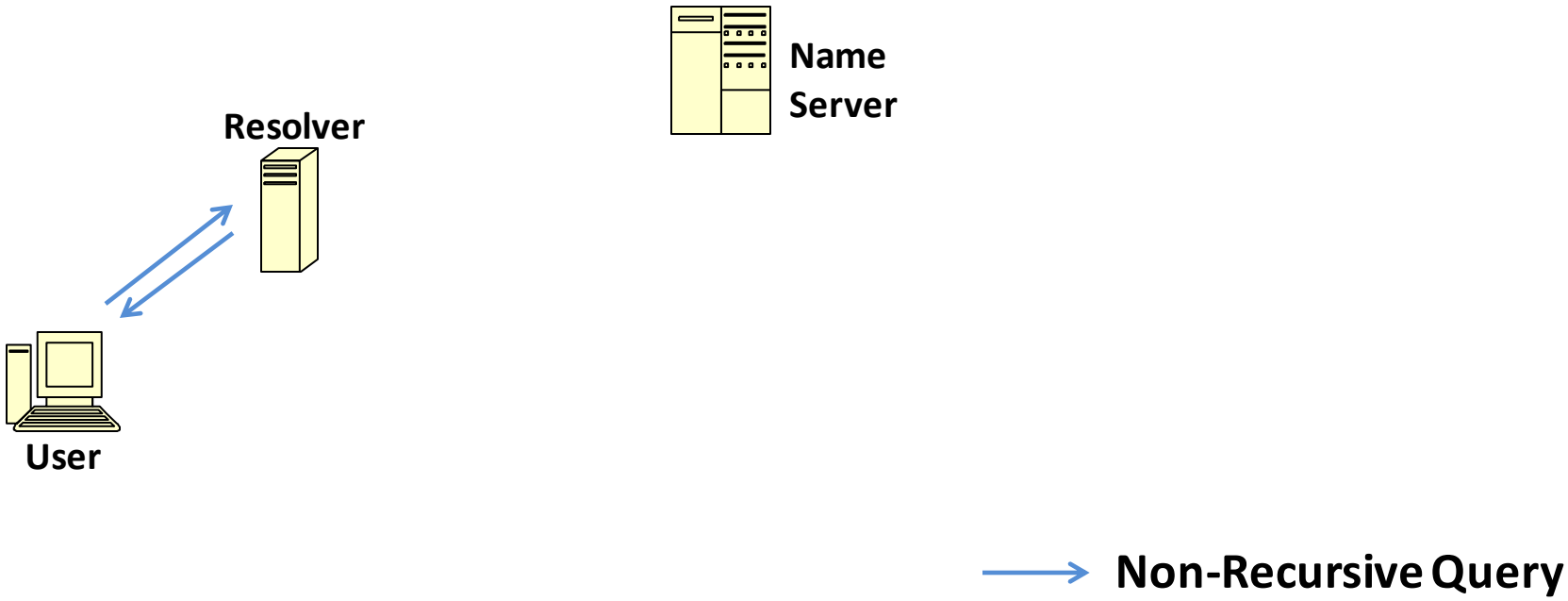
What to measure

- What is the actual effect of replications?
  - Efficient enough?
  - Uneven QoS improved?
- We need a technical survey all around the world

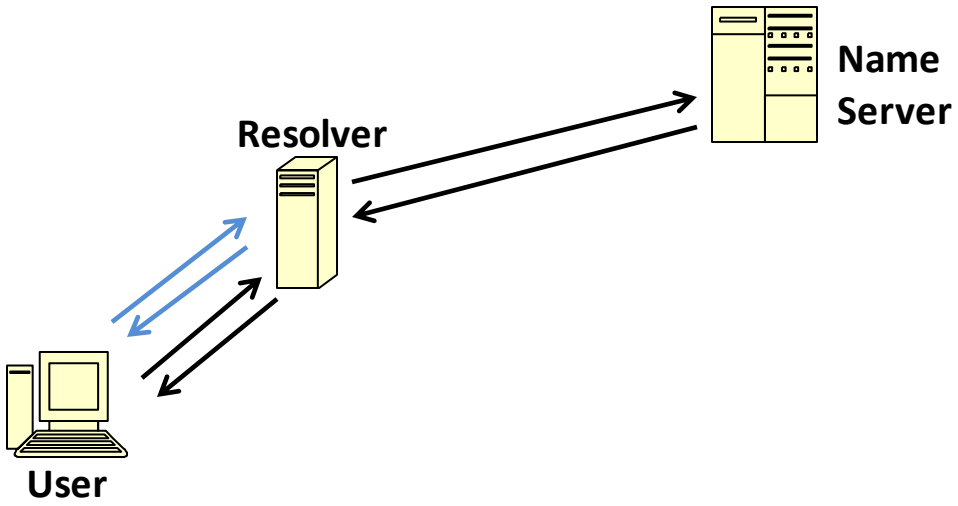
How to measure : using resolvers



How to measure : using resolvers



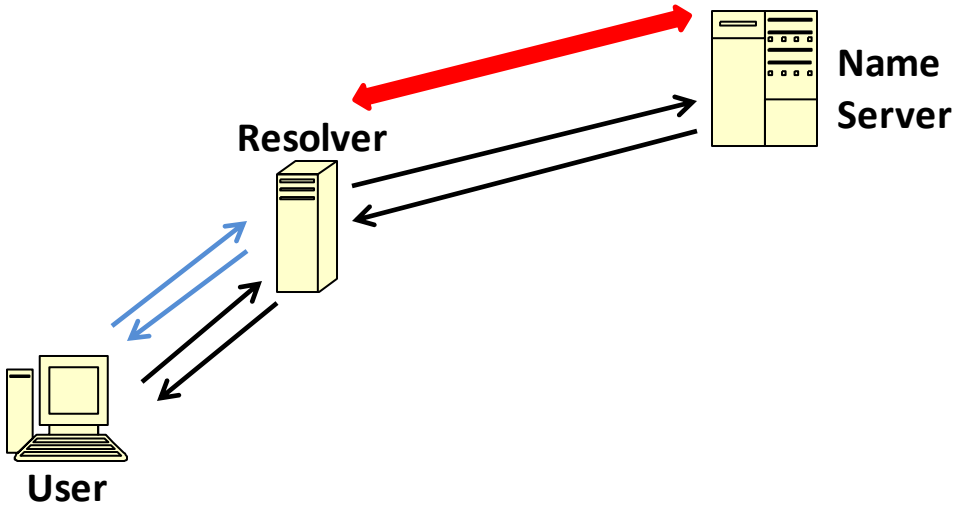
How to measure : using resolvers



Non-Recursive Query

Recursive Query

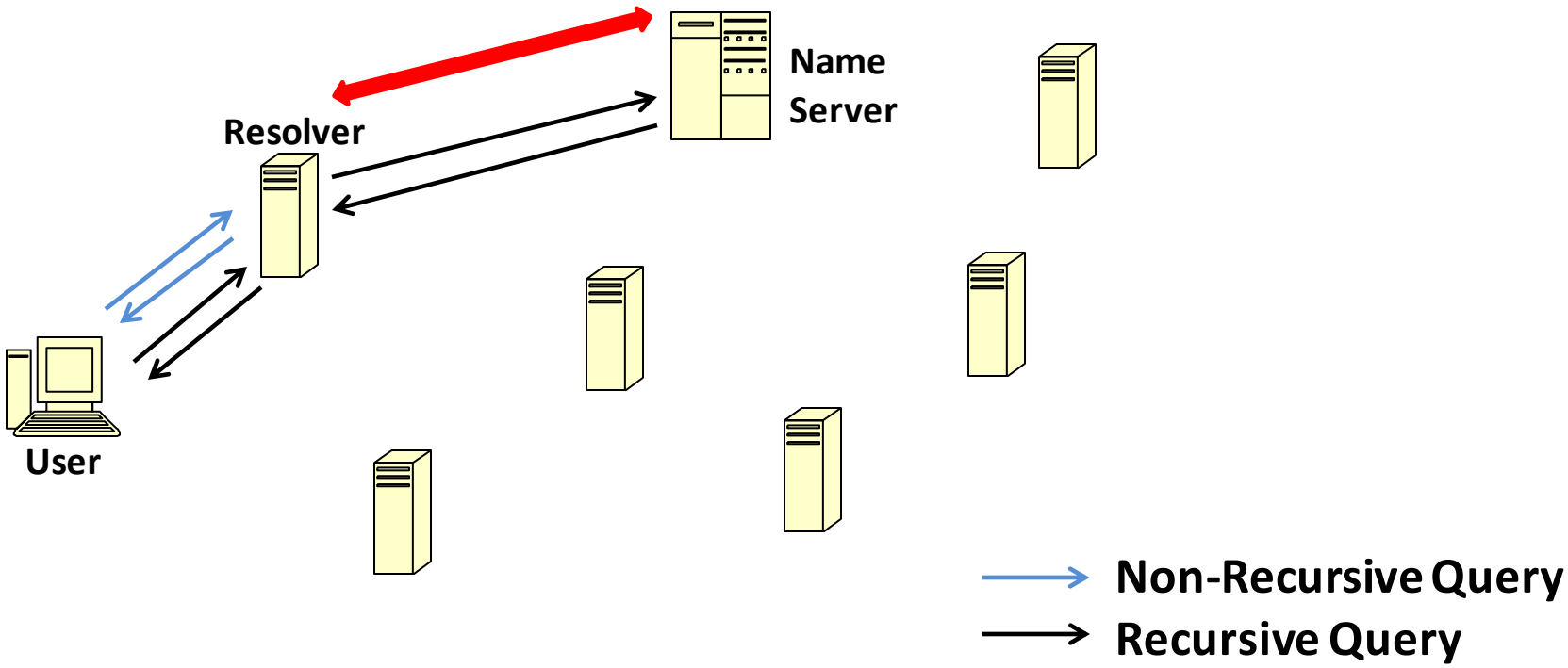
How to measure : using resolvers



Non-Recursive Query

Recursive Query

How to measure : using resolvers



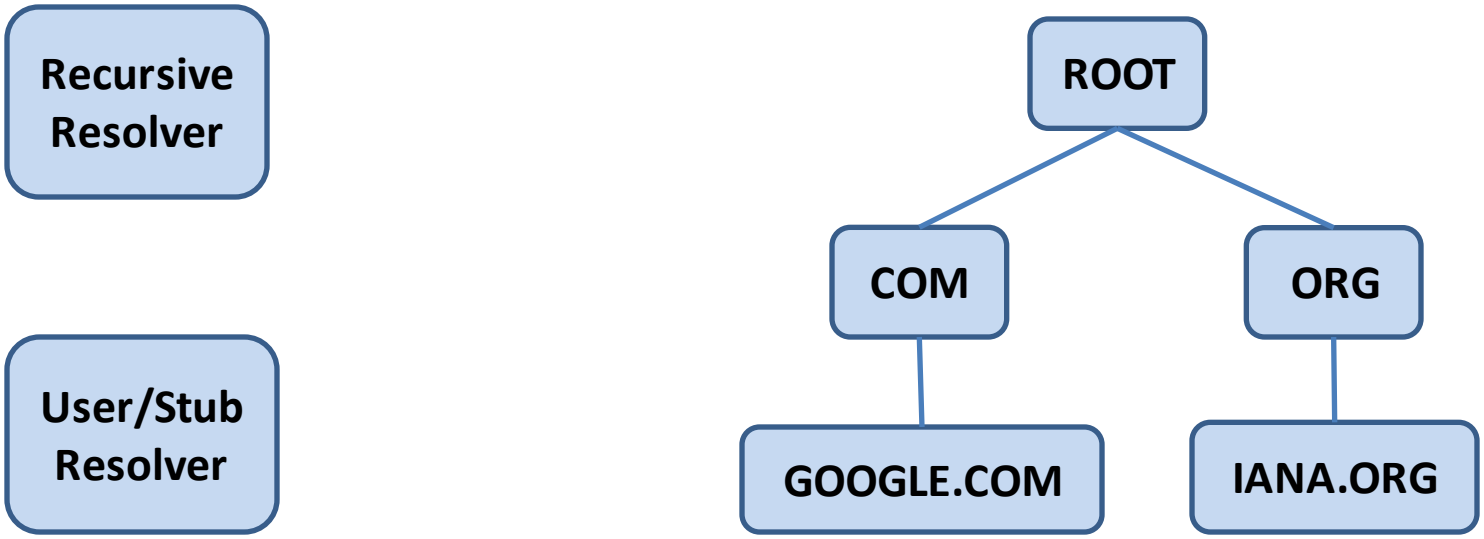
- Advantage
  - No need for direct control of vantage points, thus easy to scale up

Method: Collecting Open Resolvers

Continent	# of countries	# of ASes	# of resolvers	% of total
Europe	45	2821	7169	36.59
North America	25	1837	5525	28.20
Asia	40	940	6056	30.91
South America	11	173	426	2.17
Oceania	7	131	248	1.27
Africa	26	77	149	0.76
Unknown	-	-	20	0.10
Total	154	5979	19593	100.00

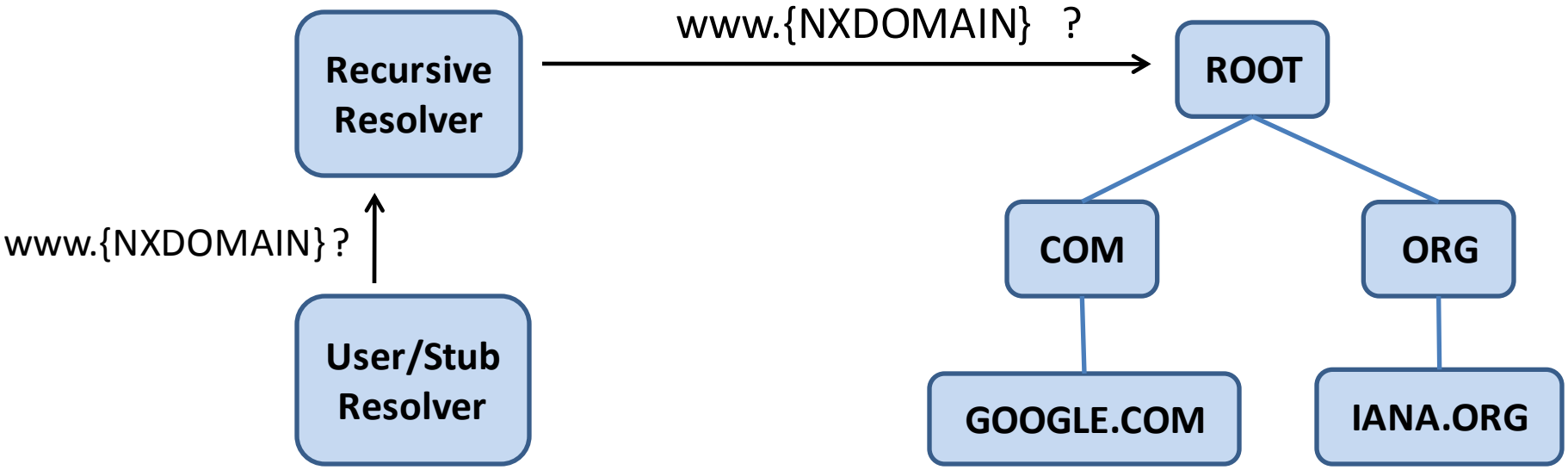
- 19593 open resolvers
  - Query log from an authority name server (42%)
  - Authority servers of Alexa top 1M sites (42%)
  - Help from other researchers (16%)
  - Exclude forwarders

Method: NXDOMAIN-Query



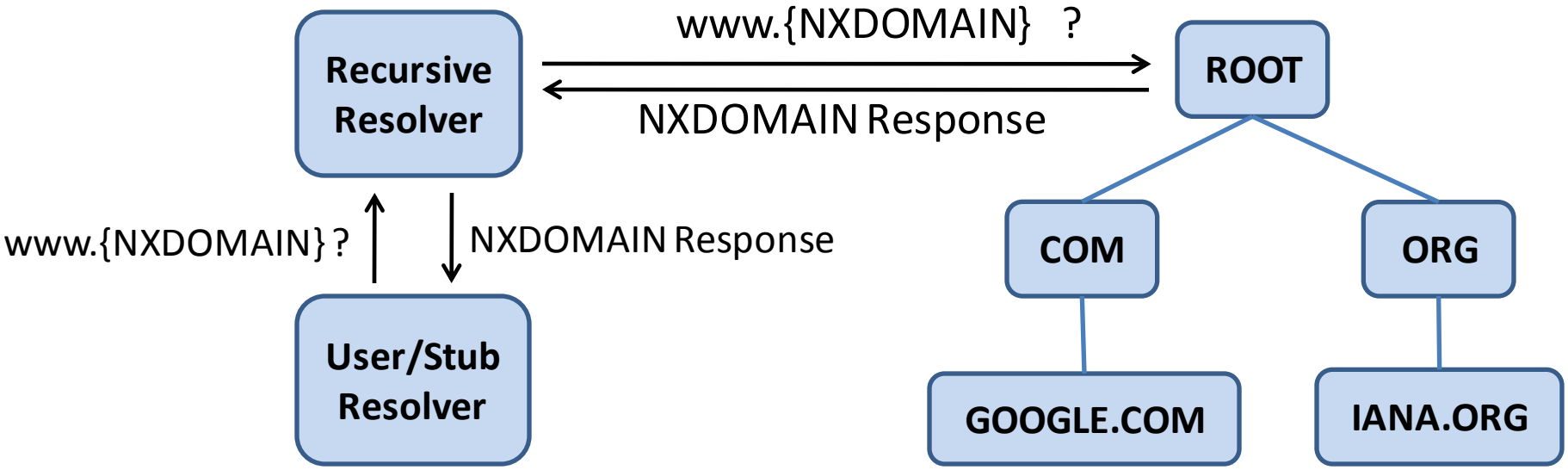
- Force a resolver to stop at a specific domain level
  - www.{NXDOMAIN}: latency to root
  - www.{NXDOMAIN}.com : latency to .com TLD
    - Don't forget to cache .com name server first

Method: NXDOMAIN-Query



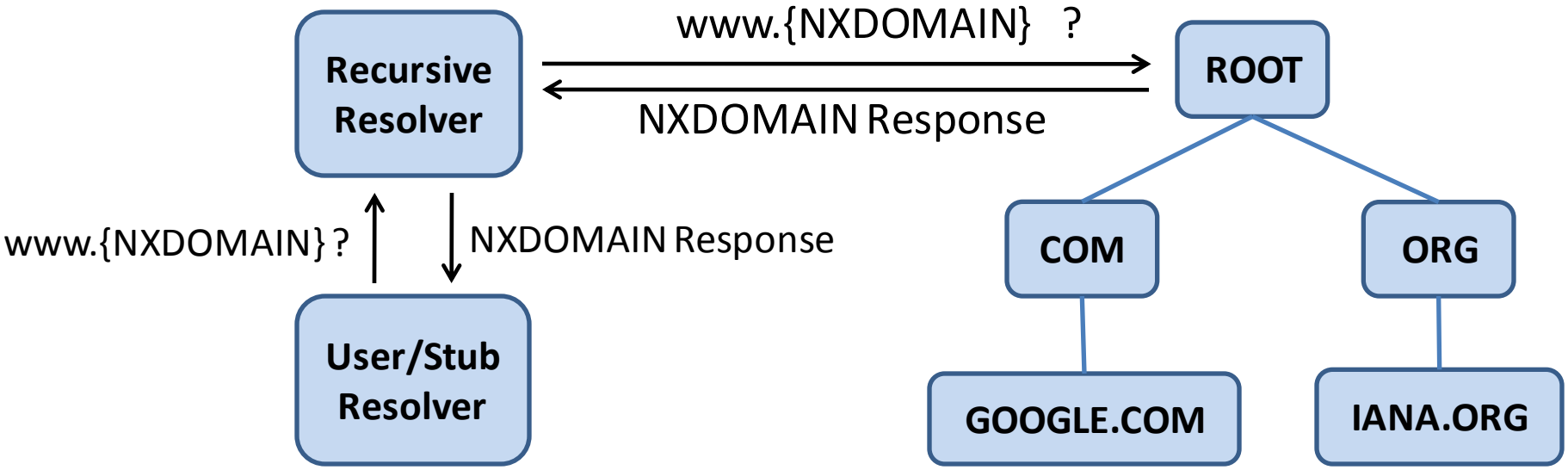
- Force a resolver to stop at a specific domain level
  - www.{NXDOMAIN}: latency to root
  - www.{NXDOMAIN}.com : latency to .com TLD
    - Don't forget to cache .com name server first

Method: NXDOMAIN-Query



- Force a resolver to stop at a specific domain level
  - www.{NXDOMAIN}: latency to root
  - www.{NXDOMAIN}.com : latency to .com TLD
    - Don't forget to cache .com name server first

Method: NXDOMAIN-Query



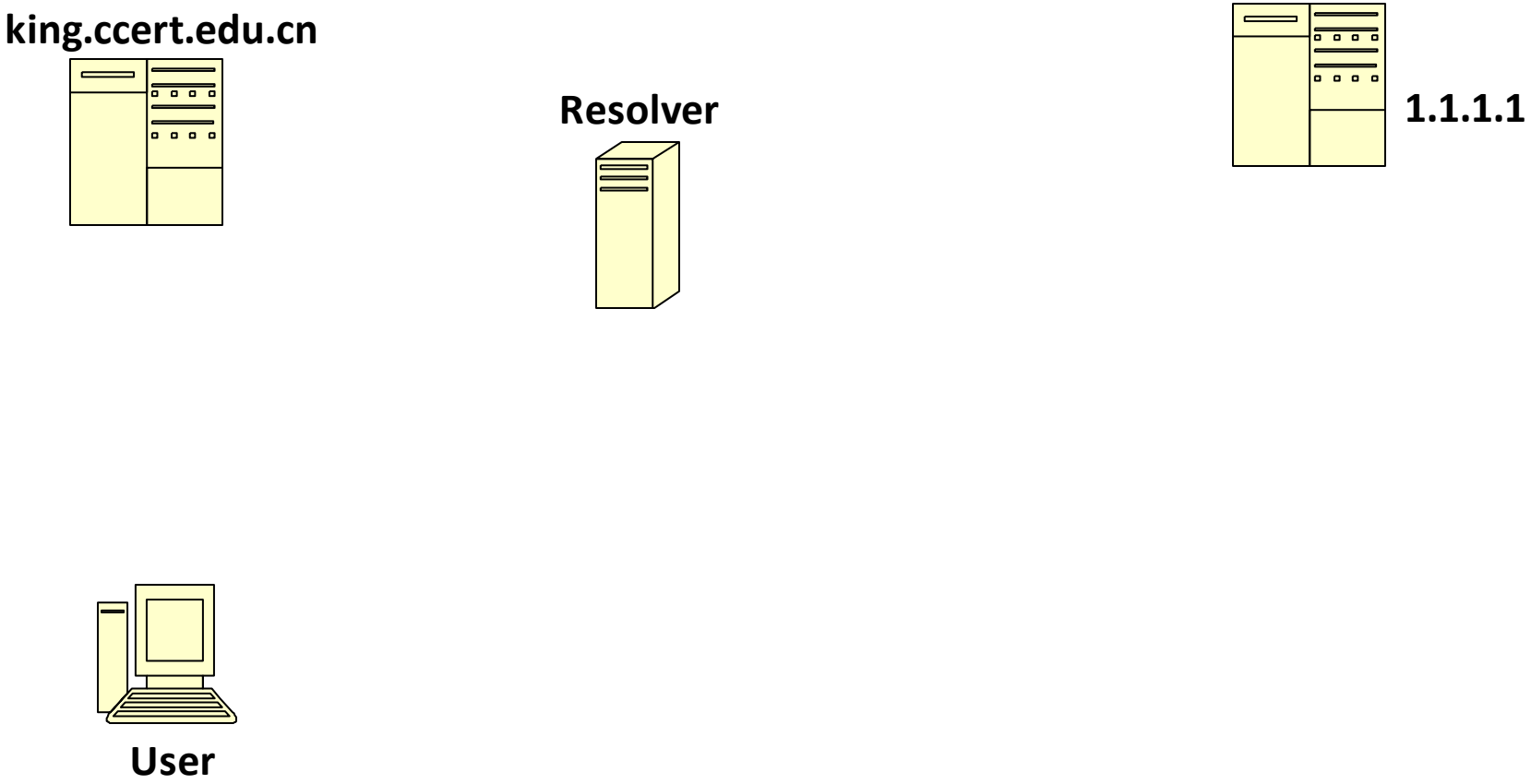
- Force a resolver to stop at a specific domain level
  - www.{NXDOMAIN}: latency to root
  - www.{NXDOMAIN}.com : latency to .com TLD
    - Don't forget to cache .com name server first
- Advantage && Limitation
  - Not affected by the cache
  - Observe latency to a domain rather than a specific server

Method: The King Technique

- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server

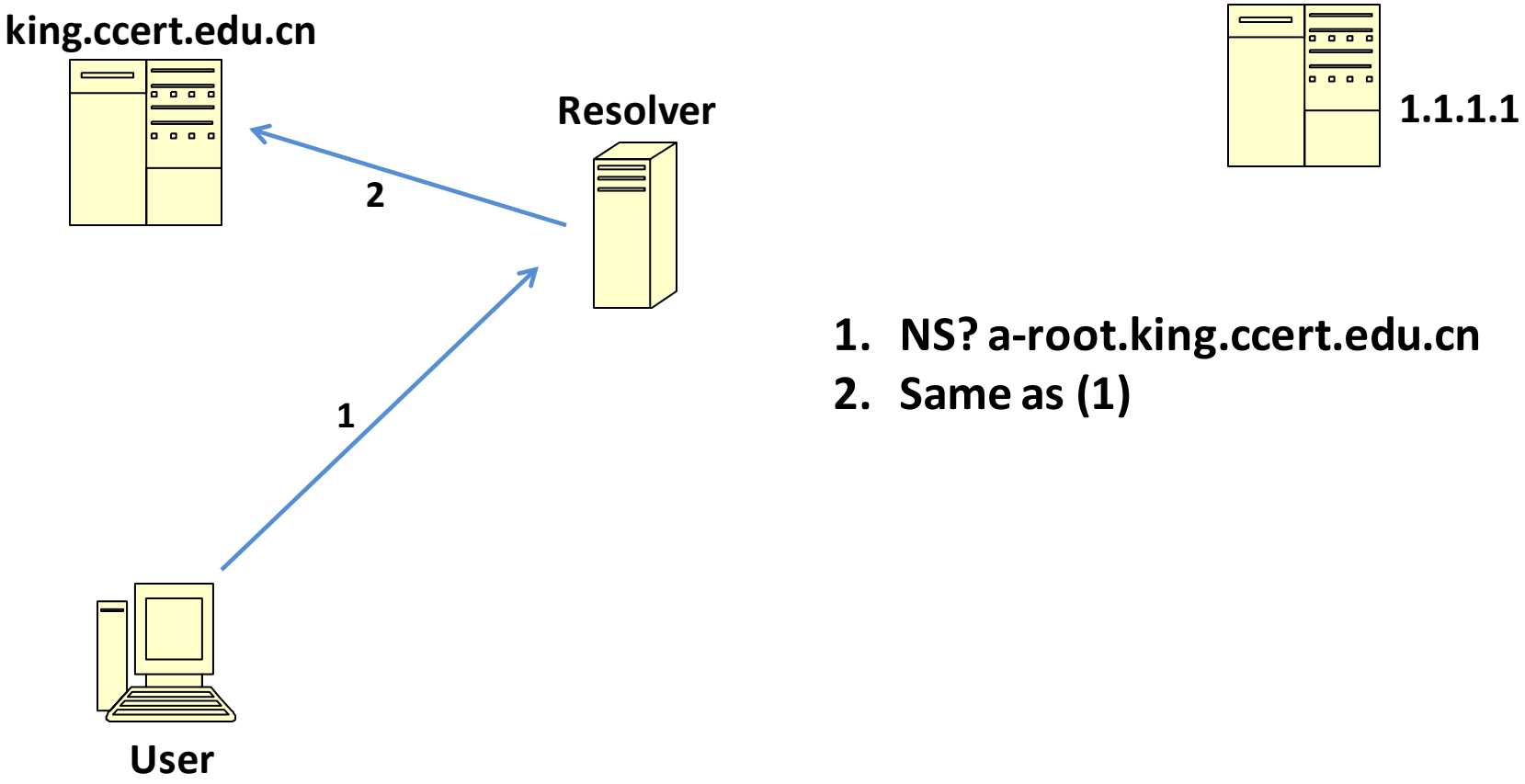
Method: The King Technique

- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



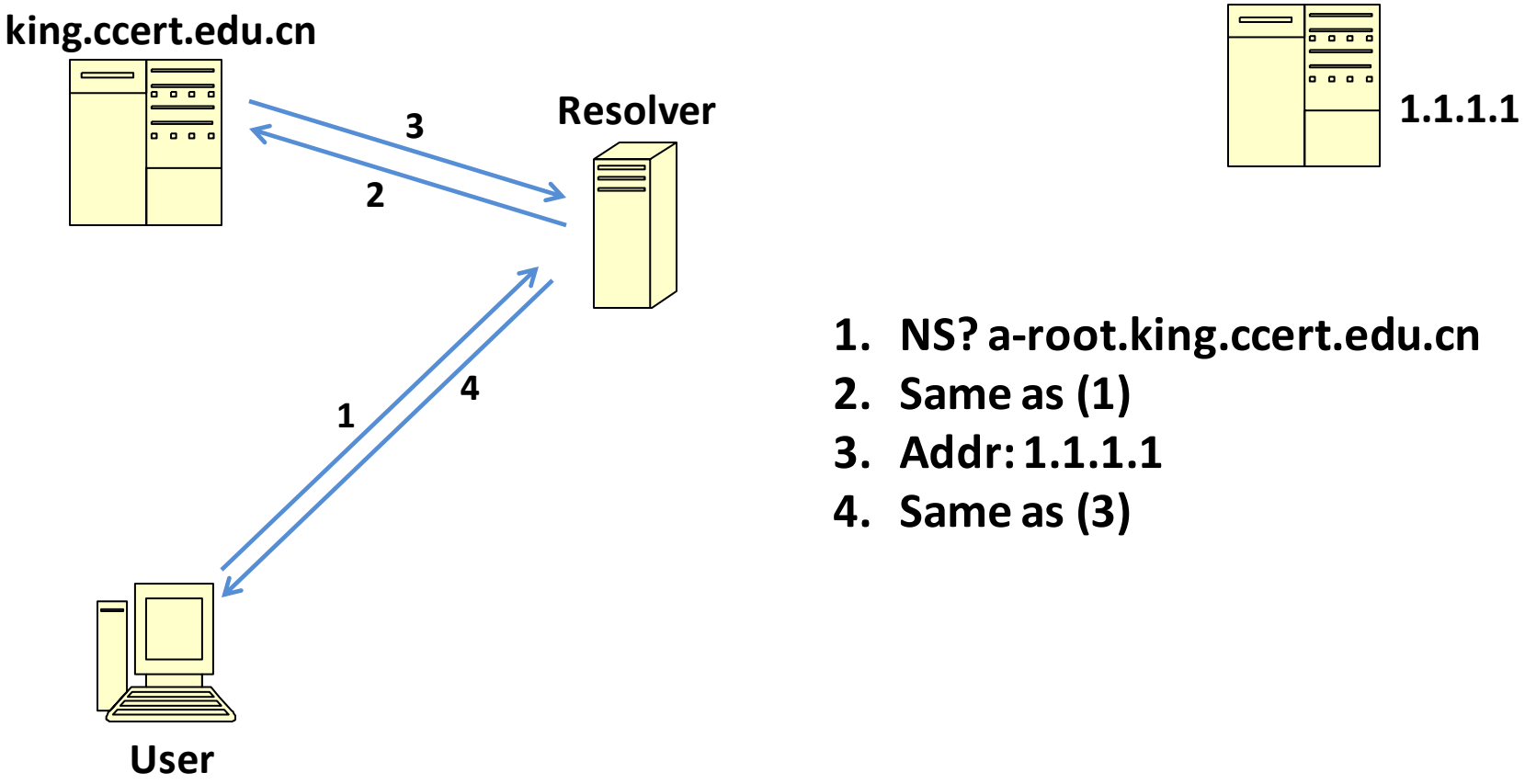
Method: The King Technique

- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



Method: The King Technique

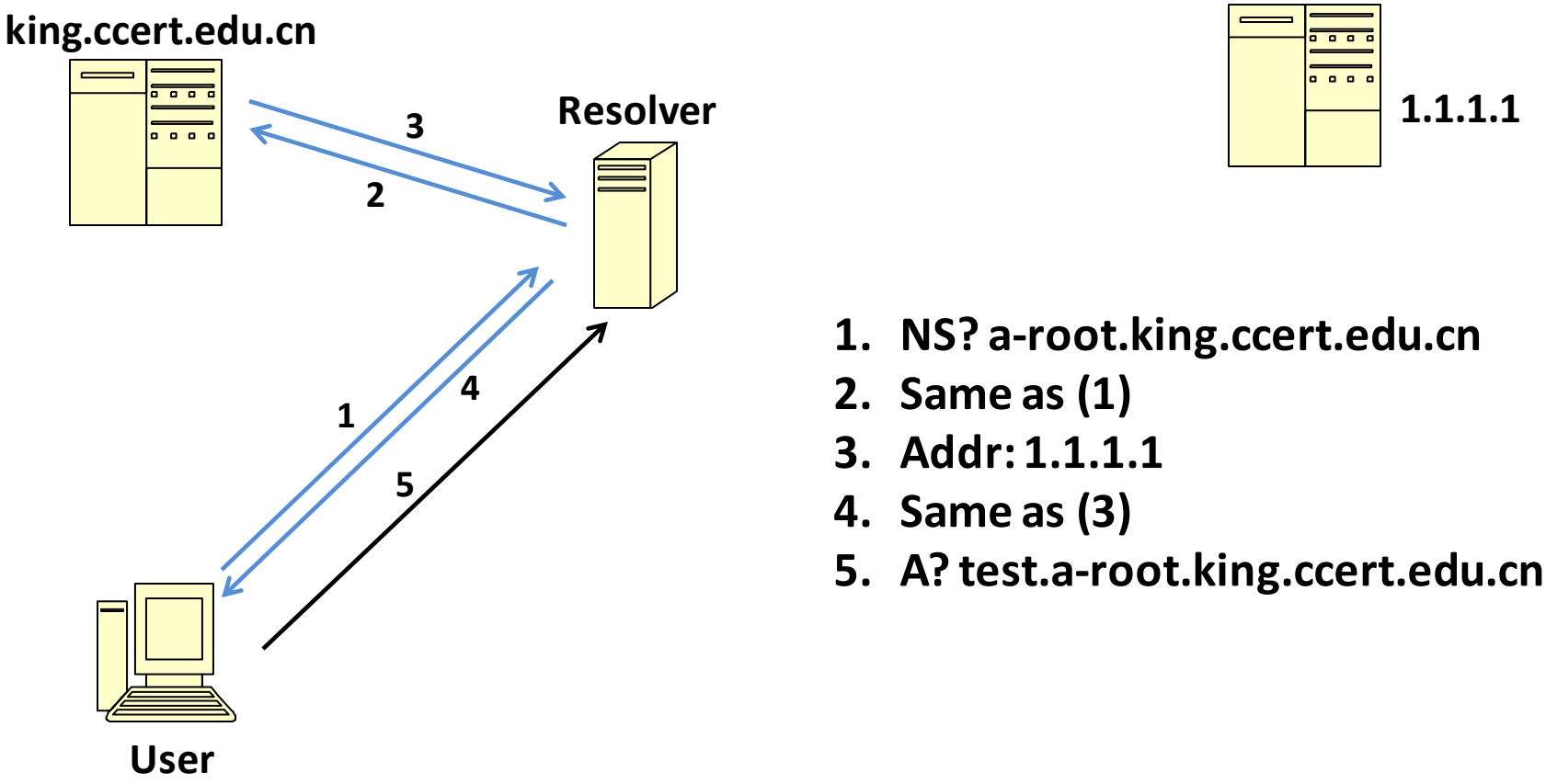
- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



1. **NS? a-root.king.ccert.edu.cn**
2. **Same as (1)**
3. **Addr: 1.1.1.1**
4. **Same as (3)**

Method: The King Technique

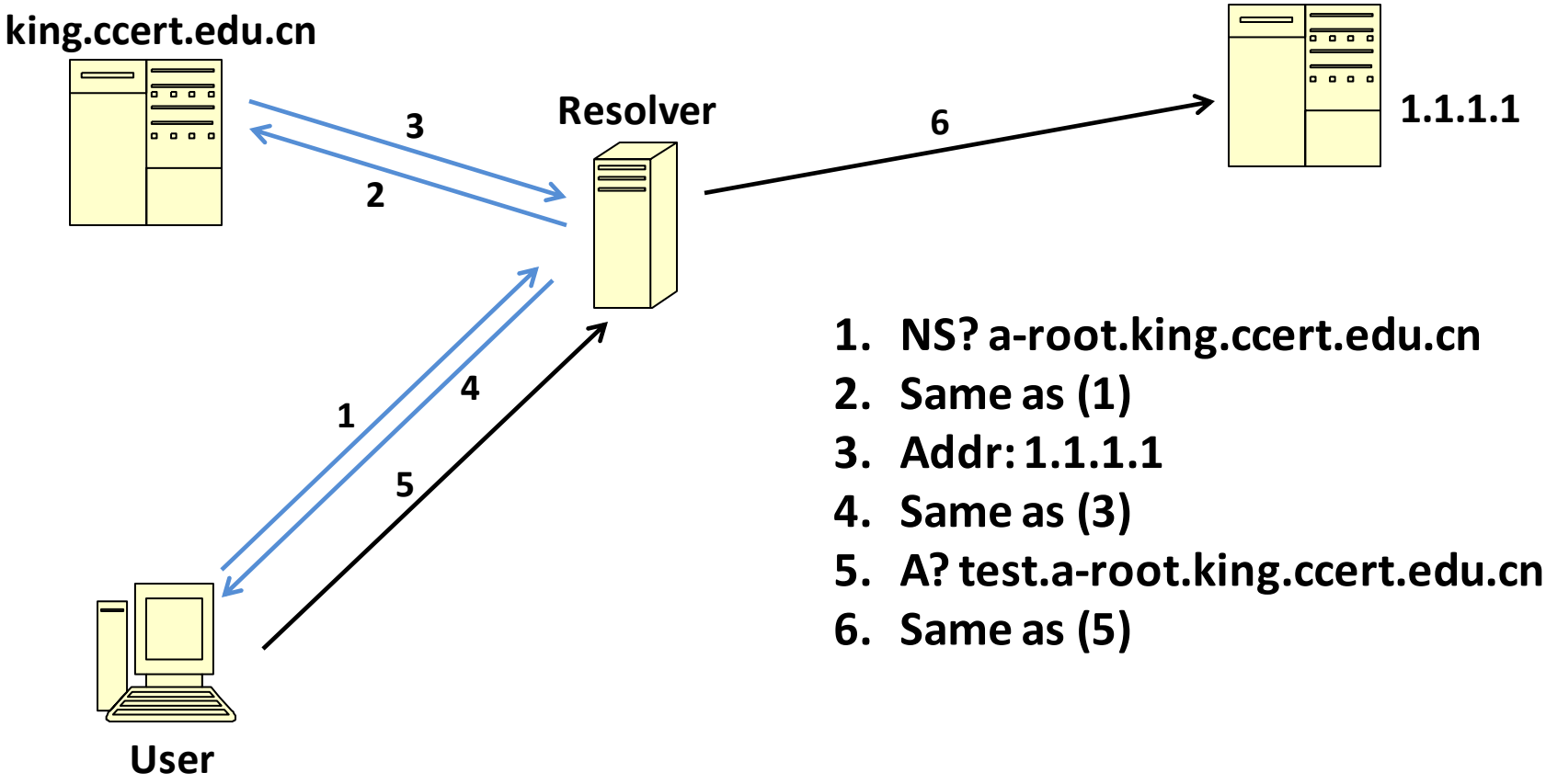
- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



1. **NS? a-root.king.ccert.edu.cn**
2. **Same as (1)**
3. **Addr: 1.1.1.1**
4. **Same as (3)**
5. **A? test.a-root.king.ccert.edu.cn**

Method: The King Technique

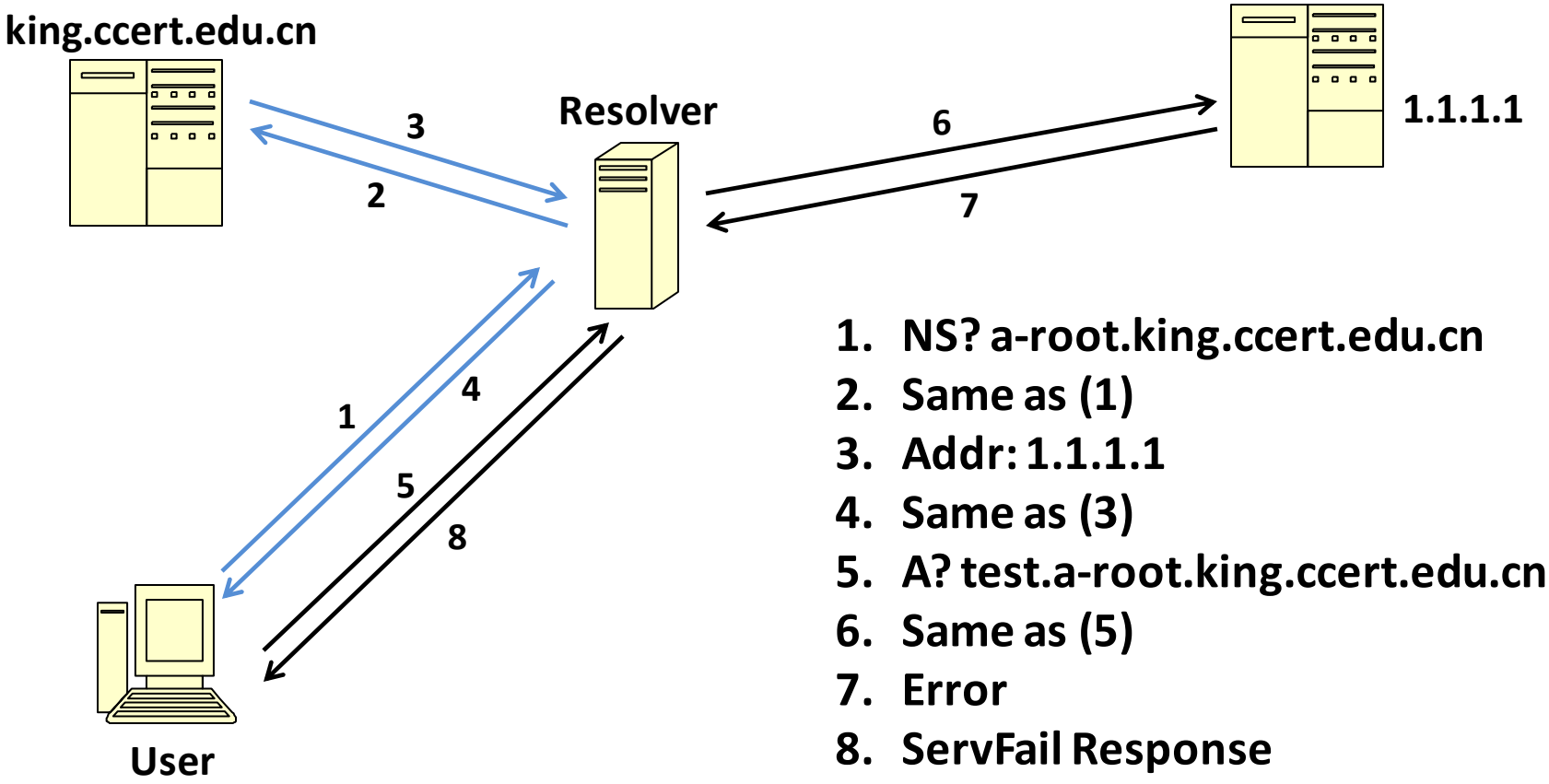
- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



1. NS? a-root.king.ccert.edu.cn
2. Same as (1)
3. Addr: 1.1.1.1
4. Same as (3)
5. A? test.a-root.king.ccert.edu.cn
6. Same as (5)

Method: The King Technique

- Measure latency from a resolver to a specific server
  - Require a controllable domain
  - Trick resolver to visit a fake name server



1. NS? a-root.king.ccert.edu.cn
2. Same as (1)
3. Addr: 1.1.1.1
4. Same as (3)
5. A? test.a-root.king.ccert.edu.cn
6. Same as (5)
7. Error
8. ServFail Response

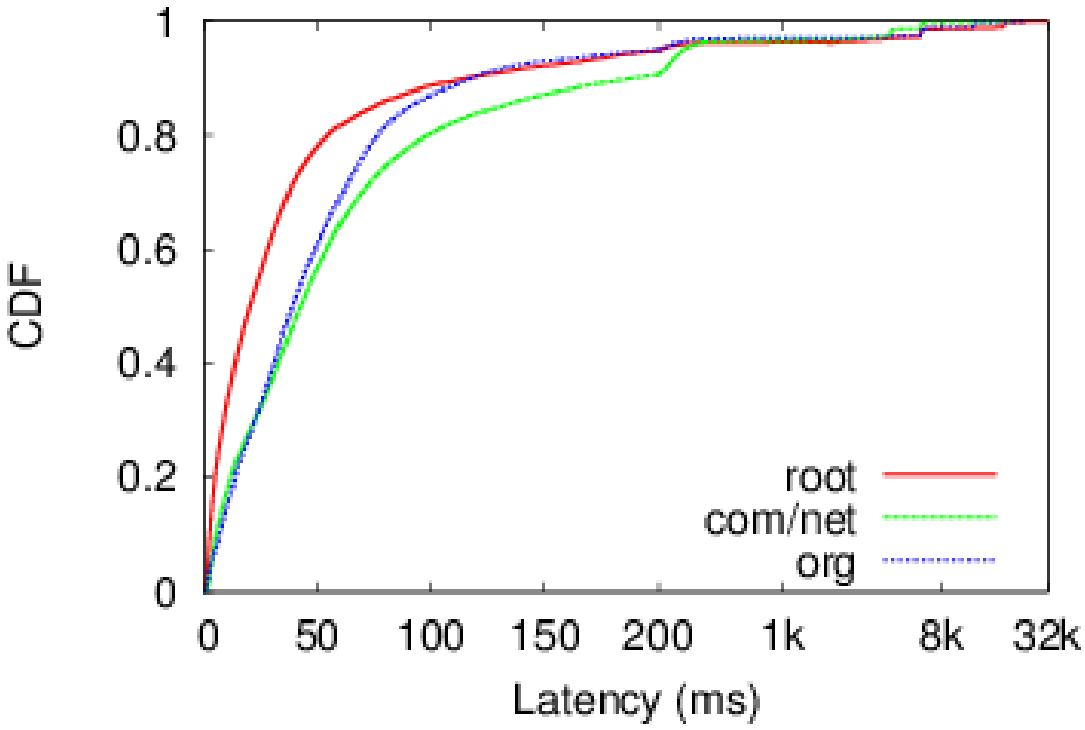
Latency of Root and TLD hierarchy

- Using NXDOMAIN-Query; root, .com/.net, .org
- 500 queries in two days; get median values

Latency of Root and TLD hierarchy

- Using NXDOMAIN-Query; root, .com/.net, .org
- 500 queries in two days; get median values

- Results
  - root (20.26ms)
  - org (39.07ms)
  - com/net (42.64ms)

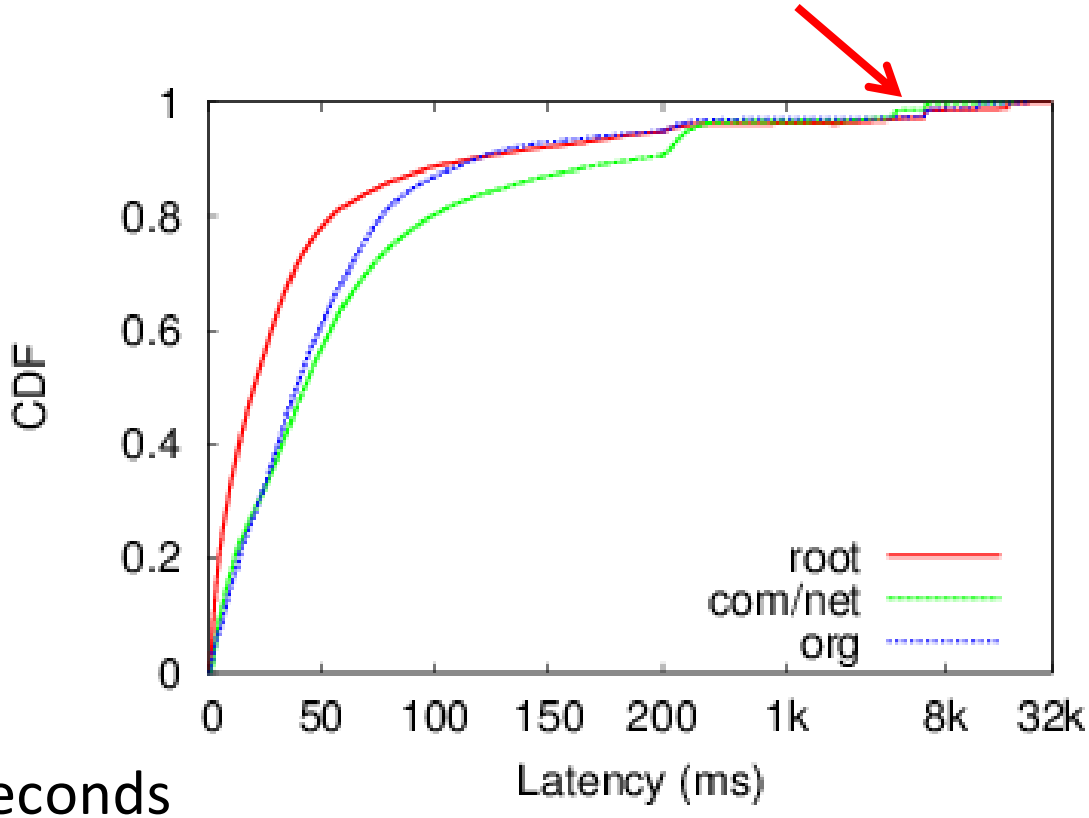


Latency of Root and TLD hierarchy

- Using NXDOMAIN-Query; root, .com/.net, .org
- 500 queries in two days; get median values

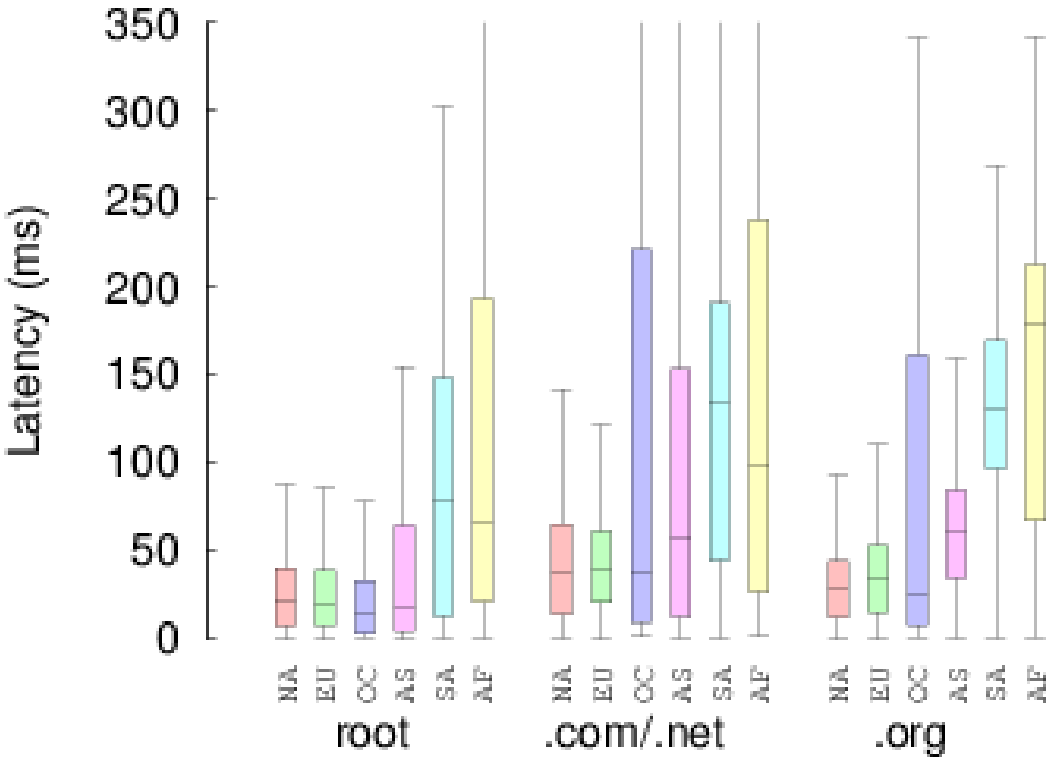
- Results
  - root (20.26ms)
  - org (39.07ms)
  - com/net (42.64ms)

  - Large query latency?
    - Around 4, 6, 12, 18 seconds



Latency of Root and TLD hierarchy

- Differences among various continents
  - Europe and North America (Best)
  - South America and Africa
    - 3 to 6 times worse
  - Oceania and Asia
    - Median values
    - Quartile values

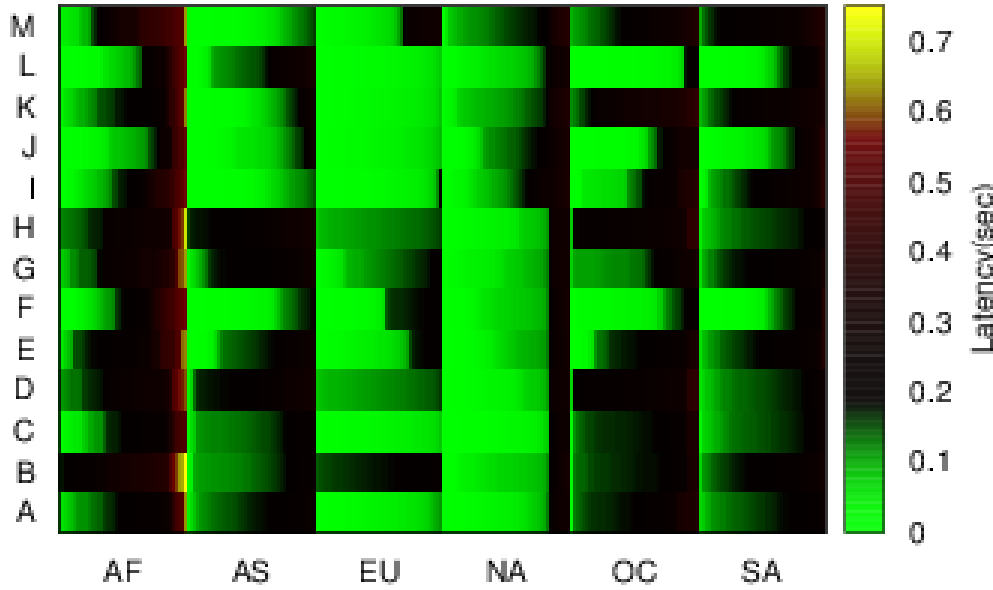


Latency of 13 root servers

- Using King technique
- 300 queries in two days; get median values

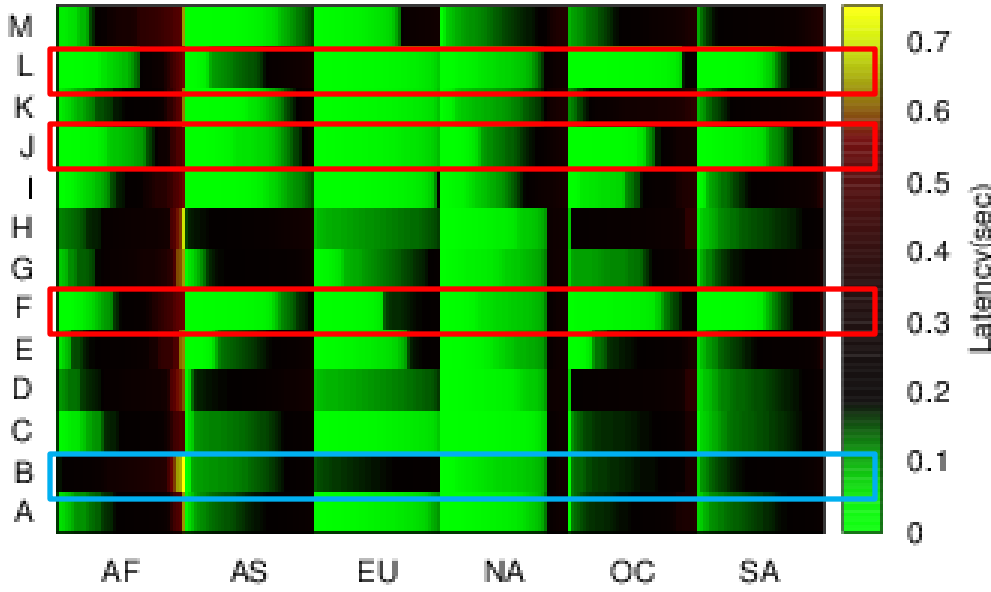
Latency of 13 root servers

- Using King technique
- 300 queries in two days; get median values



Latency of 13 root servers

- Using King technique
- 300 queries in two days; get median values
- Differences for roots
  - Best: F, J, L
    (< 200ms for continents)
  - Worst: B
    (> 300ms except NA)

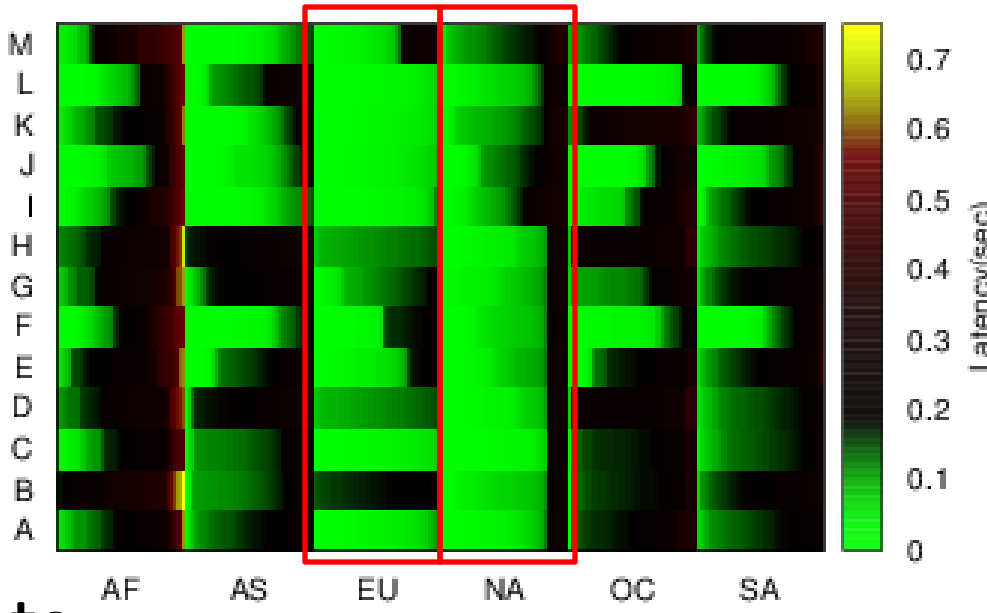


Latency of 13 root servers

- Using King technique
- 300 queries in two days; get median values
- Differences for roots
  - Best: F, J, L

    (< 200ms for continents)
  - Worst: B

    (> 300ms except NA)
- Differences for continents
  - Best: Europe & North America
  - Poor: Africa, Oceania, South America

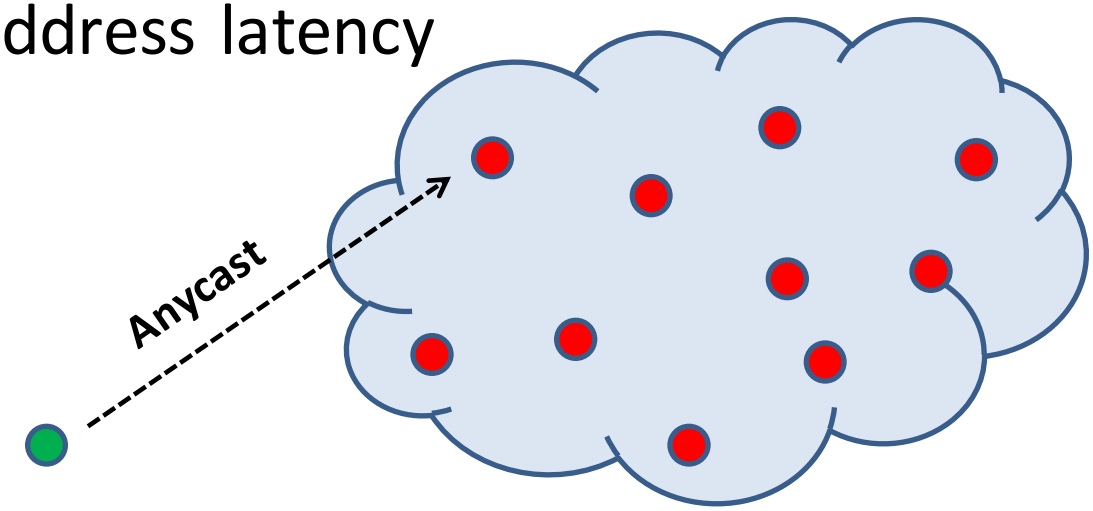


Proximity of root anycast

- What is proximity of anycast?
  - Evaluate the effect of anycast
  - Difference between anycast address latency and the minimum unicast address latency

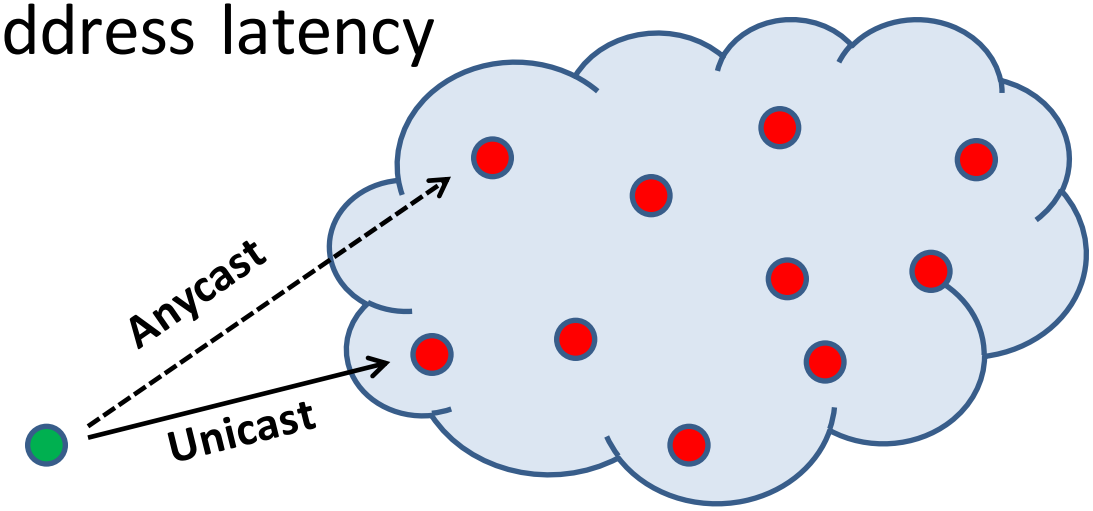
Proximity of root anycast

- What is proximity of anycast?
  - Evaluate the effect of anycast
  - Difference between anycast address latency and the minimum unicast address latency



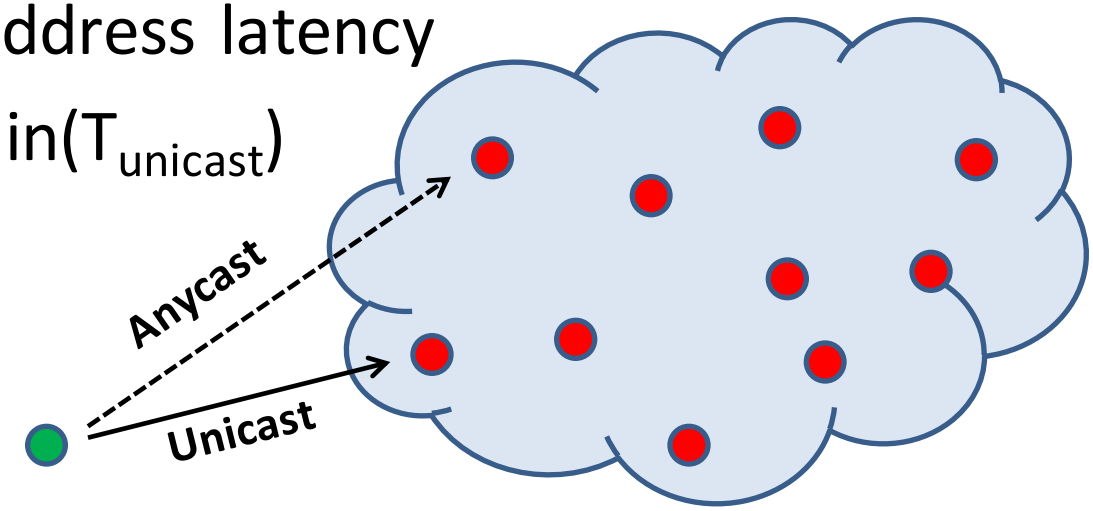
Proximity of root anycast

- What is proximity of anycast?
  - Evaluate the effect of anycast
  - Difference between anycast address latency and the minimum unicast address latency



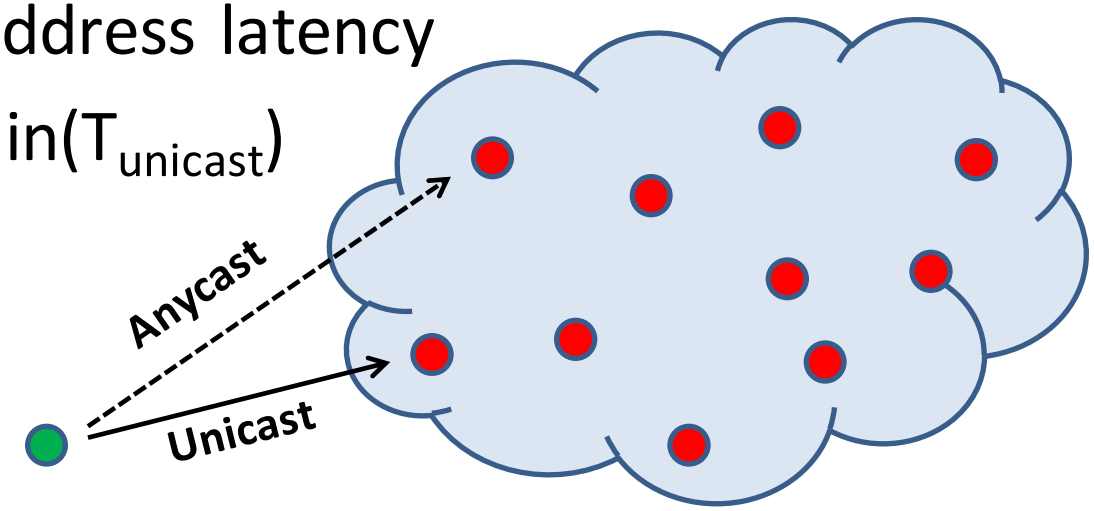
Proximity of root anycast

- What is proximity of anycast?
  - Evaluate the effect of anycast
  - Difference between anycast address latency and the minimum unicast address latency
  - $T_{proximity}=T_{anycast}-\min(T_{unicast})$



Proximity of root anycast

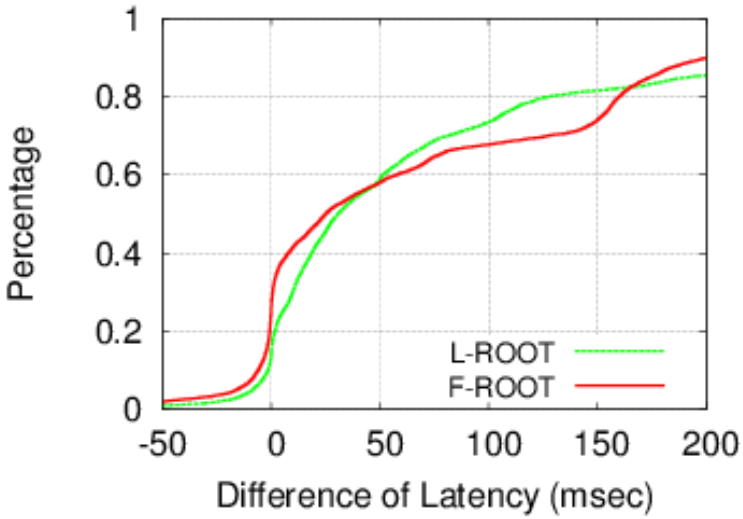
- What is proximity of anycast?
  - Evaluate the effect of anycast
  - Difference between anycast address latency and the minimum unicast address latency
  - $T_{proximity}=T_{anycast} - \min(T_{unicast})$



- Use King Technique; measure F and L root
- Repeat 200 times in 2 days; get the median values

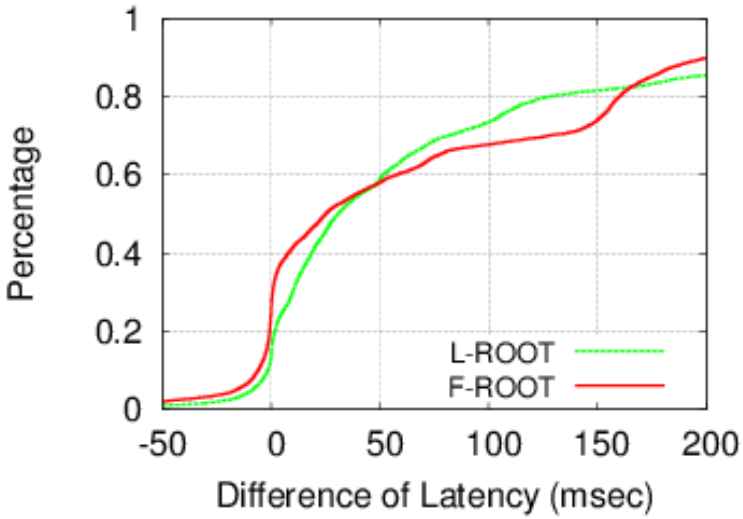
Proximity of root anycast

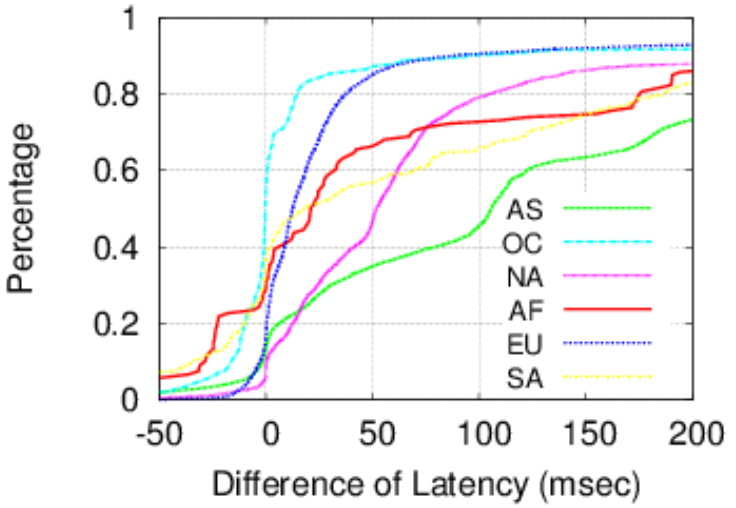
- F root && L root
  - 40% resolvers, $T_{proximity}$ > 50ms
    - Due to routing policy or hierarchical deployment
  - 2%, 1% for F and L, $T_{proximity}$ < -30ms
    - Errors in results, different routing paths, missing some unicast nodes



Proximity of root anycast

- F root && L root
  - 40% resolvers, $T_{proximity}$ > 50ms
    - Due to routing policy or hierarchical deployment
  - 2%, 1% for F and L, $T_{proximity}$ < -30ms
    - Errors in results, different routing paths, missing some unicast nodes
- L root Proximity in continents
  - Best: Oceania, Europe
  - Worst: Asia (65%, > 50ms)





Analyzing large latency

- Totally 664 resolvers (3.2% of all) constantly show large latency (> 2s)
- Root: 6s, 18s; com/net: 4s, 6s; org: 6s, 12s

- Analysis methods:
  - fpdns: get fingerprint of resolvers
  - Set up a testing domain with 3 servers to observe resolvers behavior

The cause of large latency

- Cause 1: buggy implementation on IPv4/IPv6 dual-stack
  - Software: BIND 9.2.x
  - Root: 18s; com/net: 4s; org: 12s
  - Patch: BIND (>= 9.3)

- Cause 2: filtering of DNSSEC response
  - Software: most are BIND 9.3.x
  - root, com/net, org : 6 seconds

Conclusion

- Massive deployments of server replications improve the overall DNS performance
- Quality of DNS service is still uneven among different regions
  - More anycast instances?
  - More flexible deployment policy?
- Pay more attention to the filtering of large DNSSEC packets

Thanks!

Questions?