DNS Operational Realities

John Kristoff
jtk@depaul.edu
jtk@cymru.com

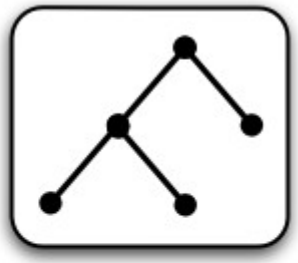
One of two critical subsystems

Practically all Internet hosts participate directly in the naming system (DNS) as a client, server or both. DNS communications may be intercepted, mangled or unreliable, but we cannot practically use the Internet without it. This can be good, bad or interesting depending on your perspective.

Fundamental DNS Components

end users
stub resolvers

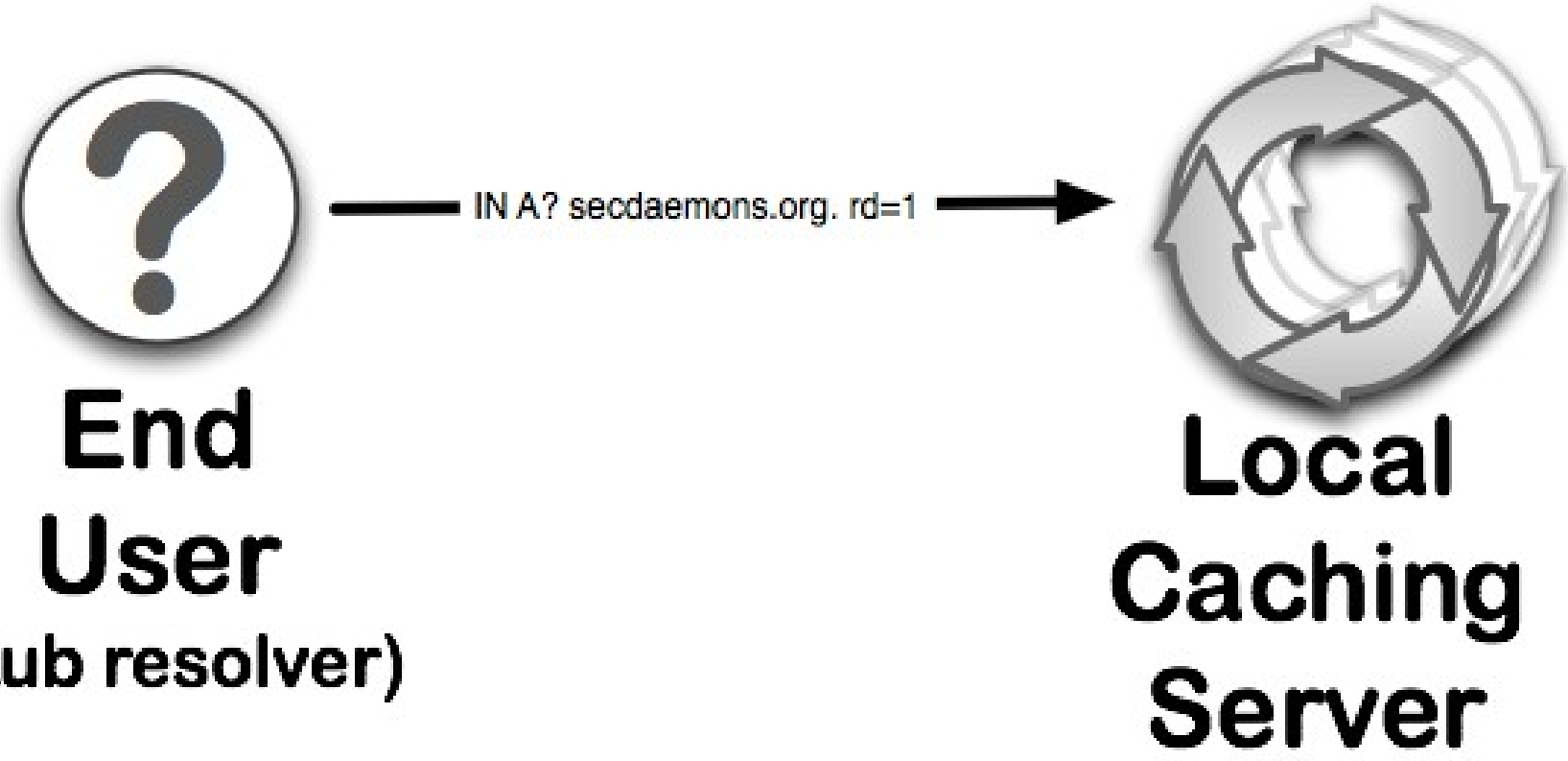
caching servers
full resolvers
forwarders



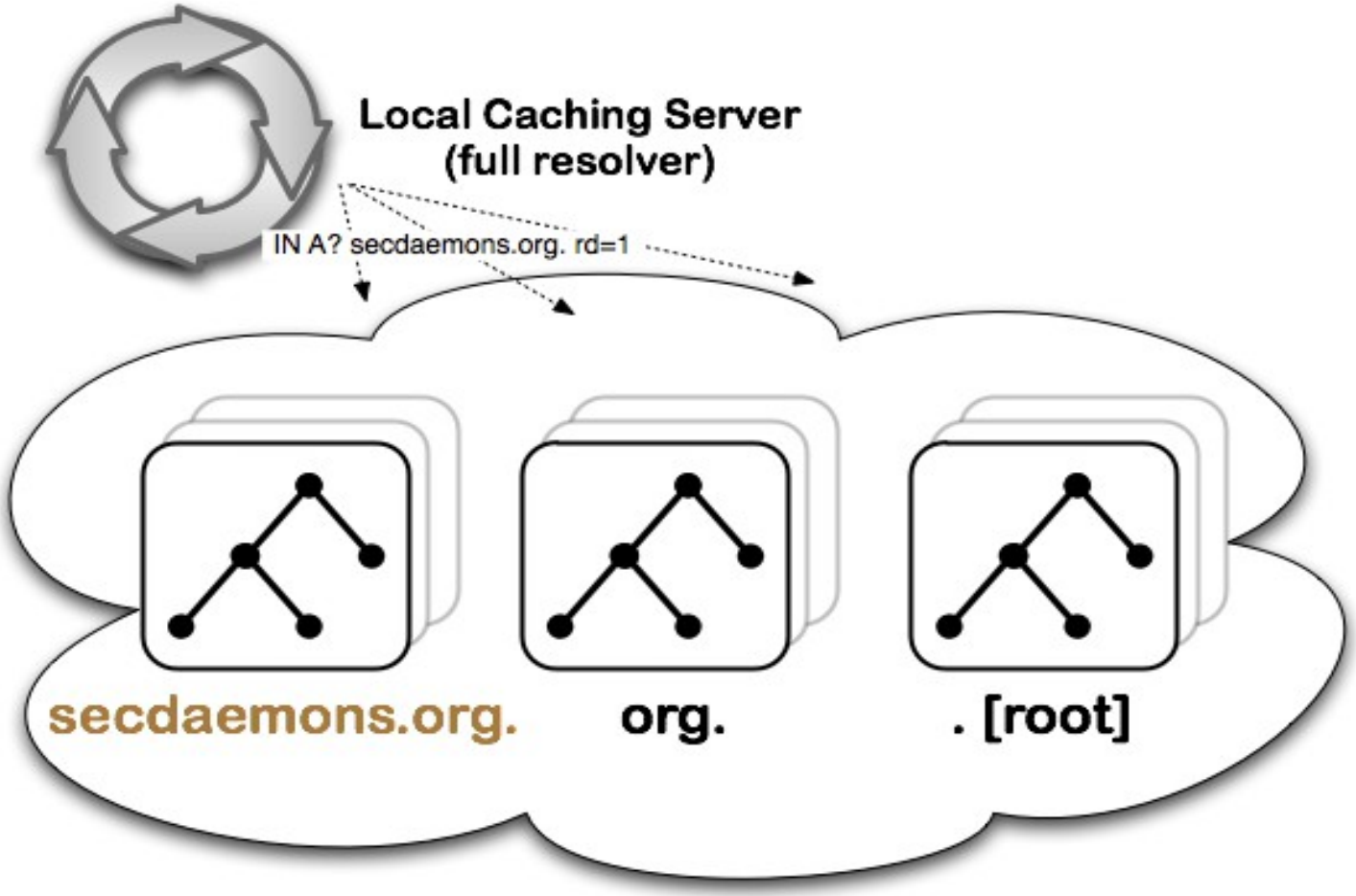
DNS name space
authoritative servers

What is the IPv4 address for secdaemons.org?

Do all the work for me (recursion desired).



1. Check cache, supply answer if available, or
2. Follow delegation from most specific cached parent, or
3. Start at root if cache is empty.



Let's assume cache is empty, and all it knows about is [.] root.*

A.root-servers.net.

...

M.root-servers.net.

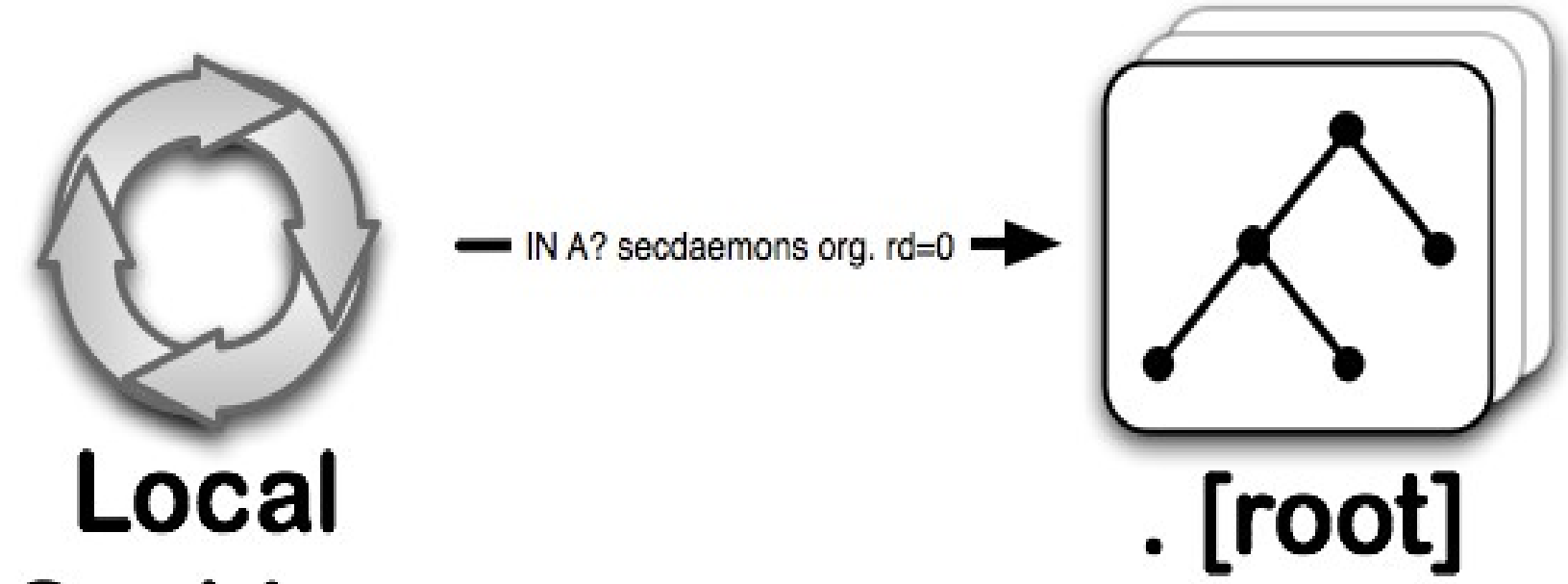
***Do you see why a reliable and trustworthy root is so important?**

What are the name servers for .org?

NO!!

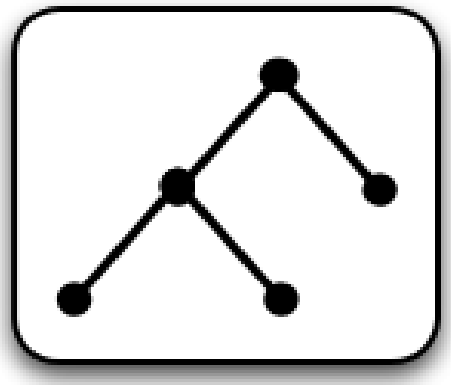
What is the IPv4 address for secdaemons.org?

I'll do the work myself (recursion disabled).



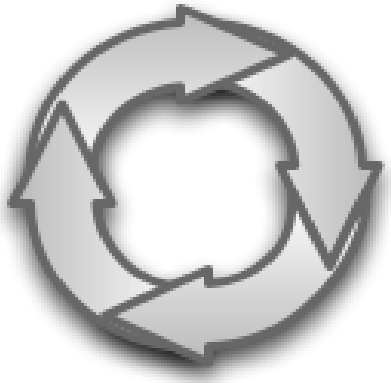
Dunno.
I refer you to .org NS RRset:

a0.org.afilias-nst.info.
a2.org.afilias-nst.info.
b0.org.afilias-nst.org.
b2.org.afilias-nst.org.
c0.org.afilias-nst.info.
d0.org.afilias-nst.org.



. [root]

NOERROR



Local Caching Server

Does the local caching server have something in its cache now?

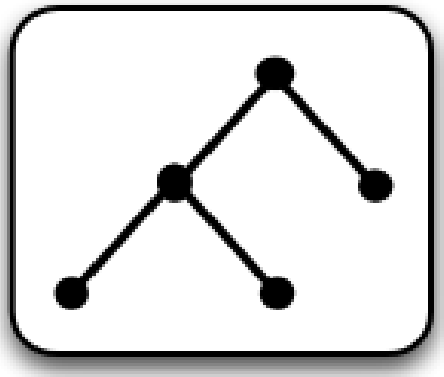
Raise your hand for yes.

Ultimately we should get here...

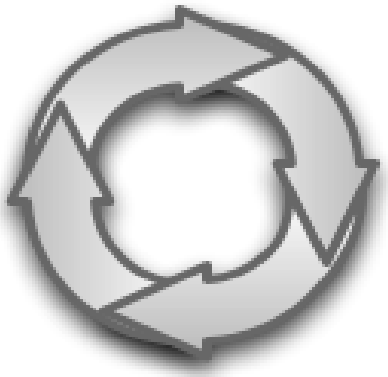
You've come to the right place. The answer RRset is:

140.192.245.240 w/ TTL=14400

I am authoritative (aa bit is set).



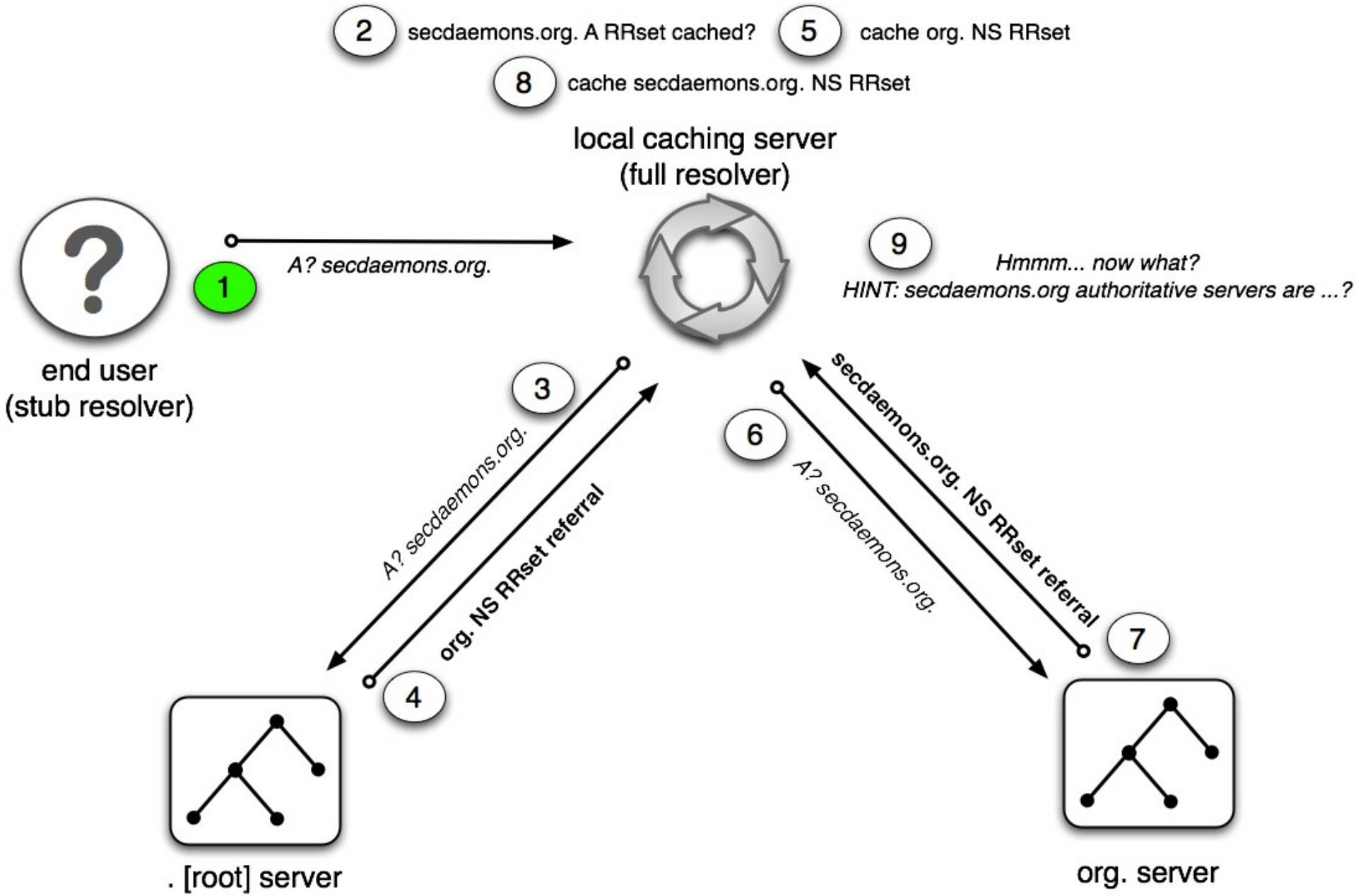
NOERROR, aa=1, 140.192.245.240

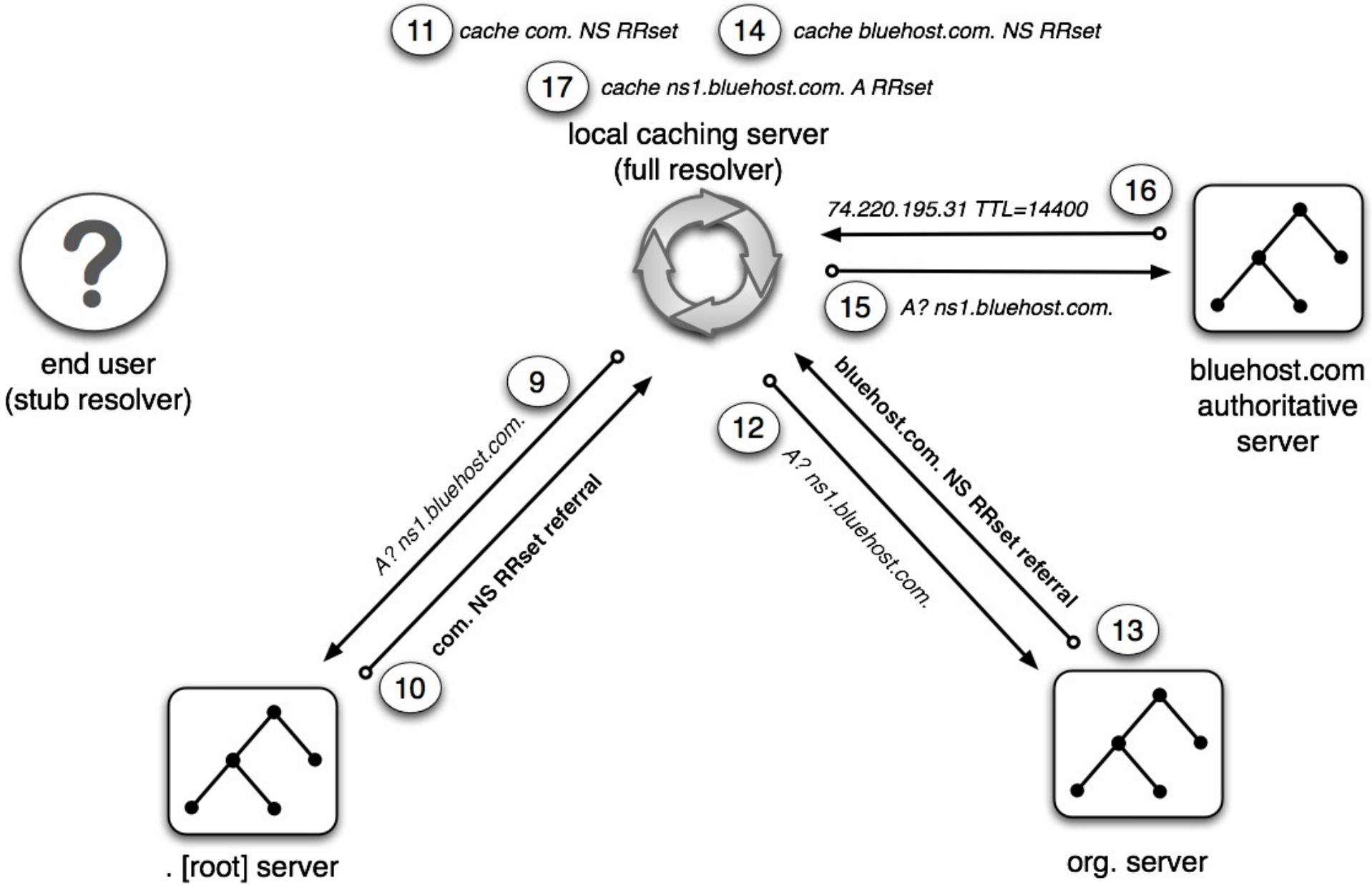


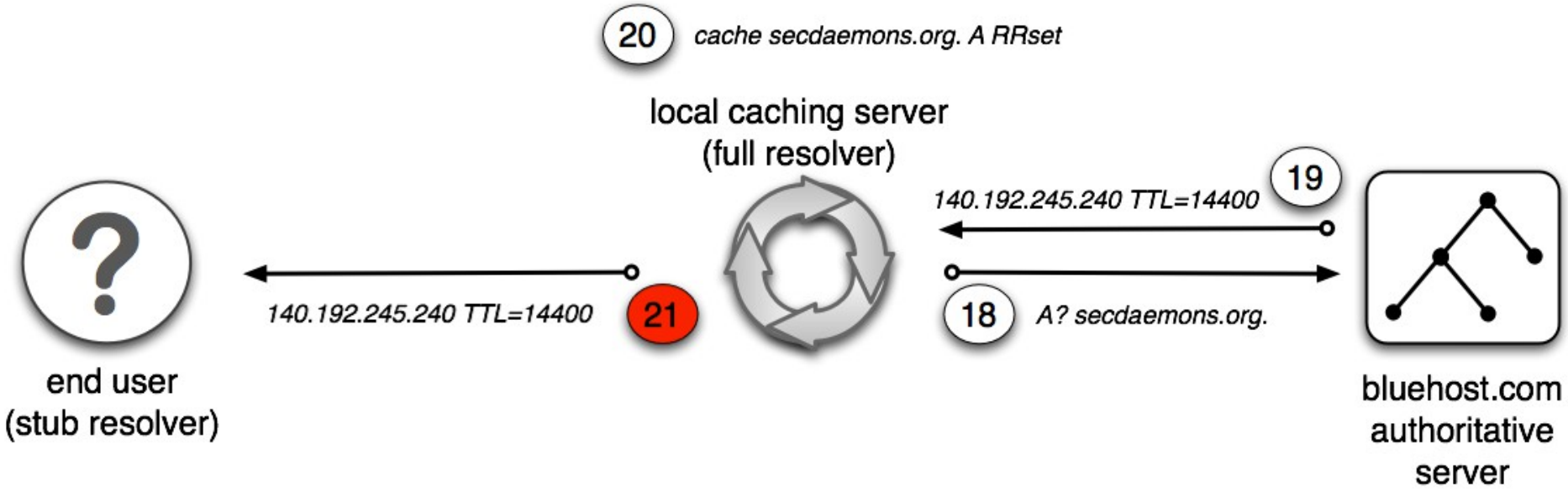
ns1.bluehost.com.
or
ns2.bluehost.com.

Local Caching Server

Big picture(s) time


2
secdaemons.org. A RRset cached?
5
cache org. NS RRset
8
cache secdaemons.org. NS RRset
local caching server
(full resolver)
9
Hmmm... now what?
HINT: secdaemons.org authoritative servers are ...?
1
A? secdaemons.org.
end user
(stub resolver)
3
A? secdaemons.org.
org. NS RRset referral
4
6
A? secdaemons.org.
secdaemons.org. NS RRset referral
7
. [root] server
org. server


11 cache com. NS RRset
14 cache bluehost.com. NS RRset
17 cache ns1.bluehost.com. A RRset
local caching server
(full resolver)
16
74.220.195.31 TTL=14400
15 A? ns1.bluehost.com.
end user
(stub resolver)
bluehost.com
authoritative
server
9
A? ns1.bluehost.com.
com. NS RRset referral
10
12
A? ns1.bluehost.com.
bluehost.com. NS RRset referral
13
. [root] server
org. server


20
cache secdaemons.org. A RRset
local caching server
(full resolver)
19
140.192.245.240 TTL=14400
18
A? secdaemons.org.
140.192.245.240 TTL=14400
21
end user
(stub resolver)
bluehost.com
authoritative
server

Now consider: cdm.depaul.edu

```
$dig +noall +norecurse +authority +additional \
     @ns1.depaul.edu cdm.depaul.edu ns

cdm.depaul.edu.         86400  IN NS ns1.cti.depaul.edu.
cdm.depaul.edu.         86400  IN NS ns2.cti.depaul.edu.
ns1.cti.depaul.edu.     2696   IN A  140.192.32.20
ns2.cti.depaul.edu.     2696   IN A  140.192.125.156
```

But ns1.cti.depaul.edu says:

```
$dig +noall +norecurse +answer \
     @ns1.cti.depaul.edu cdm.depaul.edu ns

cdm.depaul.edu.  3600 IN NS atkins.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS bach.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS cti-lpc.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS dc-colo-cti.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS ellington.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS moe.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS mozart.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS ns1.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS ns2.cti.depaul.edu.
cdm.depaul.edu.  3600 IN NS ns3.cti.depaul.edu.
cdm.depaul.edu.     0 IN NS shemp.cti.depaul.edu.
```

And additional section:

```
$dig +noall +norecurse +additional \
     @ns1.cti.depaul.edu cdm.depaul.edu ns

atkins.cti.depaul.edu.          3600 IN A  140.192.32.35
bach.cti.depaul.edu.            3600 IN A  140.192.36.3
cti-lpc.cti.depaul.edu.         1200 IN A  140.192.125.131
dc-colo-cti.cti.depaul.edu.     3600 IN A  10.128.30.2
ellington.cti.depaul.edu.       3600 IN A  140.192.32.4
moe.cti.depaul.edu.             3600 IN A  140.192.125.161
mozart.cti.depaul.edu.          3600 IN A  140.192.36.4
ns1.cti.depaul.edu.             3600 IN A  140.192.32.20
ns2.cti.depaul.edu.             3600 IN A  140.192.125.156
ns3.cti.depaul.edu.              3600 IN A  140.192.125.208
shemp.cti.depaul.edu.           1200 IN A  140.192.125.156
```

Let's test some*
(from outside of DePaul)

```
$dig +norecurse +noall +answer \
     @moe.cti.depaul.edu cdm.depaul.edu ns

;; connection timed out; no servers could be reached
```

*Results the same for: bach, mozart and ns3

Rationale for local access only?

```
$ dig @moe.cti.depaul.edu foo.edu |
      grep flags | cut -c1-53 -
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1
```

*Results the same for: bach, mozart and ns3

Unavailable from anywhere...

```
$dig +norecurse +noall +answer \
     @ns3.cti.depaul.edu cdm.depaul.edu ns

;; connection timed out; no servers could be reached
```

Also: cti-lpc

And this is a neat trick

```
$ host atkins.cti.depaul.edu
Host atkins.cti.depaul.edu not found: 3(NXDOMAIN)
```

HA!

```
$ dig @shemp.cti.depaul.edu ch txt version \
      +noall +answer +comments

;; Warning: Message parser reports malformed message packet.
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 57241
;; flags: aa rd; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; ANSWER SECTION:
version.        1476526080 IN TXT "Microsoft DNS 6.1.7601 (1DB14556)"
```

Hhmmm...

```
dig +noall @moe.cti.depaul.edu whoami.ultradns.net +answer
whoami.ultradns.net.  0  IN A  208.69.36.13

dig +noall @moe.cti.depaul.edu whoami.ultradns.net +answer
whoami.ultradns.net.  0  IN A  67.215.65.131

$ host 208.69.36.13
13.36.69.208.in-addr.arpa domain name pointer m3.chi.opendns.com.

$ host 67.215.65.131
131.65.215.67.in-addr.arpa domain name pointer hit-block.opendns.com
```

TLDs!

```
$ host 140.192.33.136
136.33.192.140.in-addr.arpa domain name pointer ppctemp.

$ host 140.192.36.33
33.36.192.140.in-addr.arpa domain name pointer colflash-iscsi.

$ host 140.192.34.32
32.34.192.140.in-addr.arpa domain name pointer cdm-e1a51bb5940.
32.34.192.140.in-addr.arpa domain name pointer msie6.
32.34.192.140.in-addr.arpa domain name pointer ectcls23.
32.34.192.140.in-addr.arpa domain name pointer animtest2.
32.34.192.140.in-addr.arpa domain name pointer 755test-tg.
```

On the bright side, it's not all bad

```
$ dig @shemp.cti.depaul.edu cdm.depaul.edu axfr +noall
; Transfer failed.
```

Then again... CDM's DNS OS for which mainstream support just ended

```
$ sudo nmap -A shemp.cti.depaul.edu
[…]
Service Info: OS: Windows; CPE: cpe:/o:microsoft:windows

Host script results:
|_smbv2-enabled: Server supports SMBv2 protocol
| smb-security-mode:
|   Account that was used for smb scripts: guest
|   User-level authentication
|   SMB Security: Challenge/response passwords supported
|_  Message signing disabled (dangerous, but default)
| smb-os-discovery:
|   OS: Windows Server 2008 R2 Standard 7601 Service Pack 1
(Windows Server 2008 R2 Standard 6.1)
|   NetBIOS computer name: SHEMP
|   Workgroup: CTI
|_  System time: 2015-01-28 17:33:58 UTC-6
```

Time permitting...
An operational BoF preview

DNS over TCP queries...

a) SHOULD (or MUST) NOT be filtered

b) MAY (or MUST) be filtered

IETF RFC 1033 (November 1987)
DOMAIN ADMINISTRATORS OPERATIONS GUIDE

No applicable mention of transport protocol requirements.

IETF RFC 1034 (November 1987)
DOMAIN NAMES – CONCEPTS AND FACILITIES

3.7 Queries

"In the Internet, queries **are** carried in UDP datagrams or over TCP connections."

IETF RFC 1035 (November 1987)
DOMAIN NAMES – IMPLEMENTATION AND SPECIFICATION

4.2 Transport

"While virtual circuits **can be** used for any DNS activity, datagrams are preferred for queries due to their lower overhead and better performance."

IETF RFC 1123 (October 1989)

Requirements for Internet Hosts -- Application and Support

6.1.3.2 Transport Protocols

"DNS resolvers and recursive servers MUST support UDP, and **SHOULD** support TCP, for sending (non-zone-transfer) queries. […] A name server MAY limit the resources it devotes to TCP queries, but it **SHOULD NOT** refuse to service a TCP query just because it would have succeeded with UDP."

IETF RFC 1536 (October 1993)

note: category = informational

Common DNS Implementation Errors and Suggested Fixes

1. Fast Retransmissions

"DNS implements the classic request-response scheme of client-server interaction. **UDP is**, therefore, the chosen protocol for communication though TCP is used for zone transfers."

IETF RFC 2136 (April 1997)

Dynamic Updates in the Domain Name System (DNS UPDATE)

2.1 – Transport Issues

"An update transaction **may be** carried in a UDP datagram, if the request fits, or in a TCP connection (at the discretion of the requestor)."

7.8

"It is possible for a UDP response to be lost in transit and for a request to be retried due to a timeout condition. […] For this reason,requestors who require an accurate response code **must** use TCP."

IETF RFC 2181 (July 1997)
Clarifications to the DNS Specification

9. The TC (truncated) header bit

"When a DNS client receives a reply with TC set, it **should** ignore that response, and query again, using a mechanism, such as a TCP connection, that will permit larger replies."

IETF RFC 2541 (March 1999)
DNS Security Operational Considerations

4. Public/Private Key Size Considerations

"[...] larger keys increase the size of the KEY and SIG RRs. This increases the chance of DNS UDP packet overflow and the **possible necessity** for using higher overhead TCP in responses."

note: ultimately obsoleted by two later RFCs, neither of which mention TCP

IETF RFC 2671 (August 1999)
Extension Mechanisms for DNS (EDNS0)

4.5.4

Somewhat out of context, but note the underlying capability provided that suggests larger UDP messages using EDNS0

“[...] is considered **preferrable** to the outright use of TCP for oversized requests, if there is any reason to suspect that the responder implements EDNS, and if a request will not fit in the default 512 payload size limit.)”

IETF RFC 4033 (March 2005)
DNS Security Introduction and Requirements

9. Name Server Considerations

“Because inclusion of these DNSSEC RRs could easily cause UDP message truncation and fallback to TCP, a security-aware name server **must also** support the EDNS "sender's UDP payload" mechanism.”

IETF RFC 5966 (August 2010)
DNS Transport over TCP – Implementation Requirements

1. Introduction

"Whilst this document makes no specific recommendations to operators of DNS servers, it **should** be noted that failure to support TCP (or the blocking of DNS over TCP at the network layer) **may** result in resolution failure and/or application-level timeouts"

Are Operators in Agreement with DNS over TCP?

- ISC KB #AA-01219 seems to conflate implementation with operational requirements

  "DNS queries using TCP (best current practice for many years, and now clarified and asserted in RFC 5966 [...]"

- In Geoff Huston's "A Question of DNS Protocols"

  "This data appears to point to a level of failure to followup from a truncated UDP response to a TCP connection of some 2.6% of clients."

- DNS over TCP drafts and discussion on the rise

  draft-ietf-dnsop-edns-tcp-keepalive-01

  draft-wouters-edns-tcp-chain-query-01

  draft-ietf-dnsop-dnssec-roadblock-avoidance-01

Is there any DNS over TCP?

- Some, but proportionally it is a very small amount
- When is it used?
  - Larger answers due to TC=1 switchover
  - Sometimes TC=1 switchover for DDoS mitigation
  - Very rarely initiated by the client

Is there a DNS over TCP danger?

- TCP-based resource consumption attacks
  - May be a preferable challenge than UDP DDoS
- Zone transfer threat probably blown out of proportion
- There is also a danger if you don't allow it
  - Some names just won't resolve
  - Stuck TCP state on resolvers trying filtered authoritative servers (reverse resource exhaustion) – see NANOG 32 talk “DNS Anomalies and Their Impacts on DNS Cache Servers”

Missing RFC?

- “DNS over TCP Operational Requirements”

Thank you

- John Kristoff
- <jtk@depaul.edu> - https://condor.depaul.edu/jkristof/
- <jtk@cymru.com> - https://www.cymru.com/jtk/
- FYI... I'm not on linkedin, twitter, facebook …
- But I shouldn't be too hard to find