Università degli Studi Roma Tre
Dipartimento di Informatica e Automazione
Computer Networks Research Group

netkit lab

dns

Version	2.2
Author(s)	G. Di Battista, M. Patrignani, M. Pizzonia, F. Ricci, M. Rimondini
E-mail	contact@netkit.org
Web	http://www.netkit.org/
Description	using the domain name system

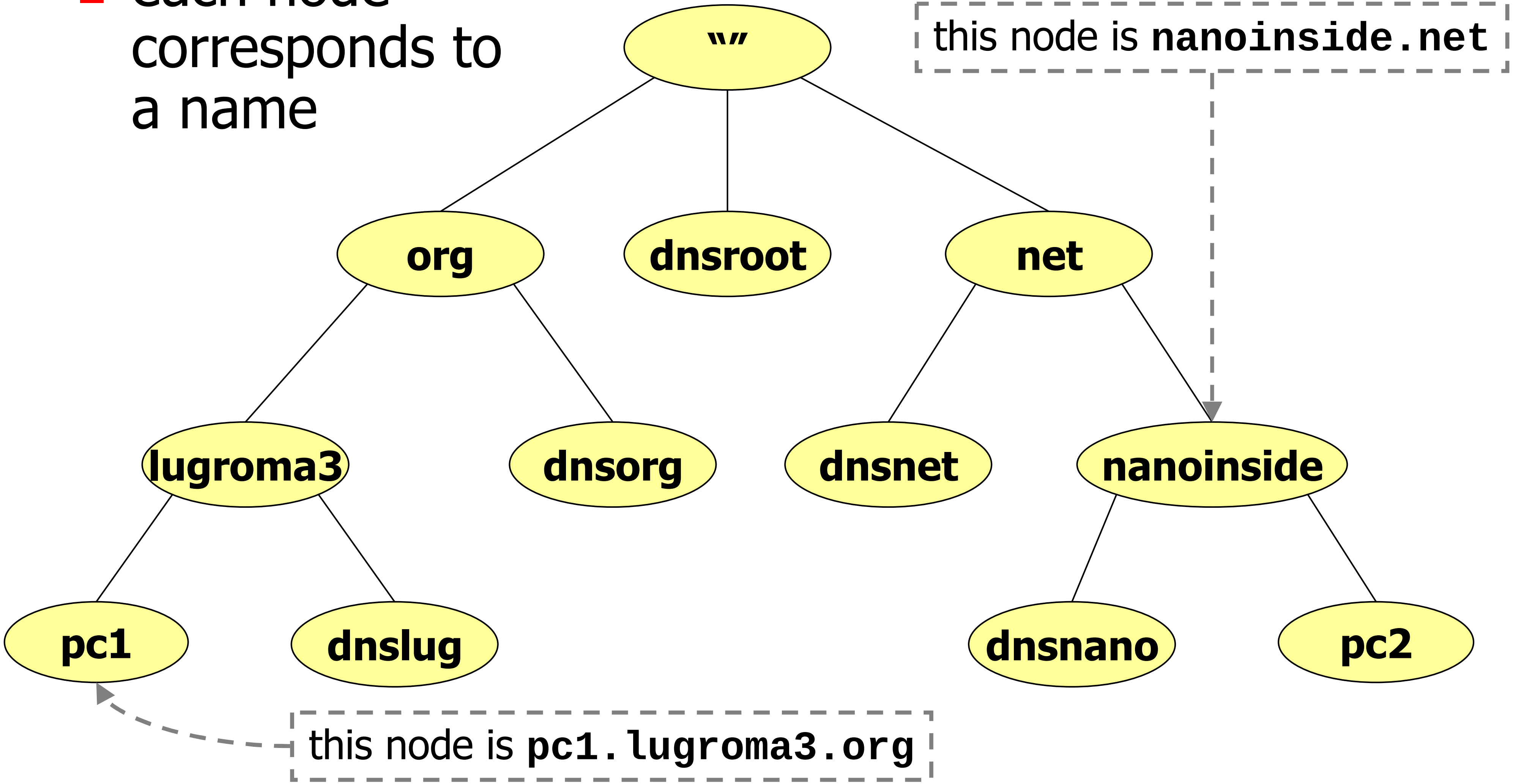
copyright notice

about the dns

- takes care of associating names with ip addresses (and more…)
- the name system is distributed over several nodes (hosts) that are hierarchically organized to form a tree
- each node in the hierarchy corresponds to a name
- a domain in the name system is a subtree
- a node in the hierarchy may be delegated to handle names for a particular zone
  - such a node is an authoritative server for that zone
- a zone is a domain which is devoid of those nodes having a different authoritative server (i.e., a tree without subtrees)

the dns name hierarchy

- each node corresponds to a name



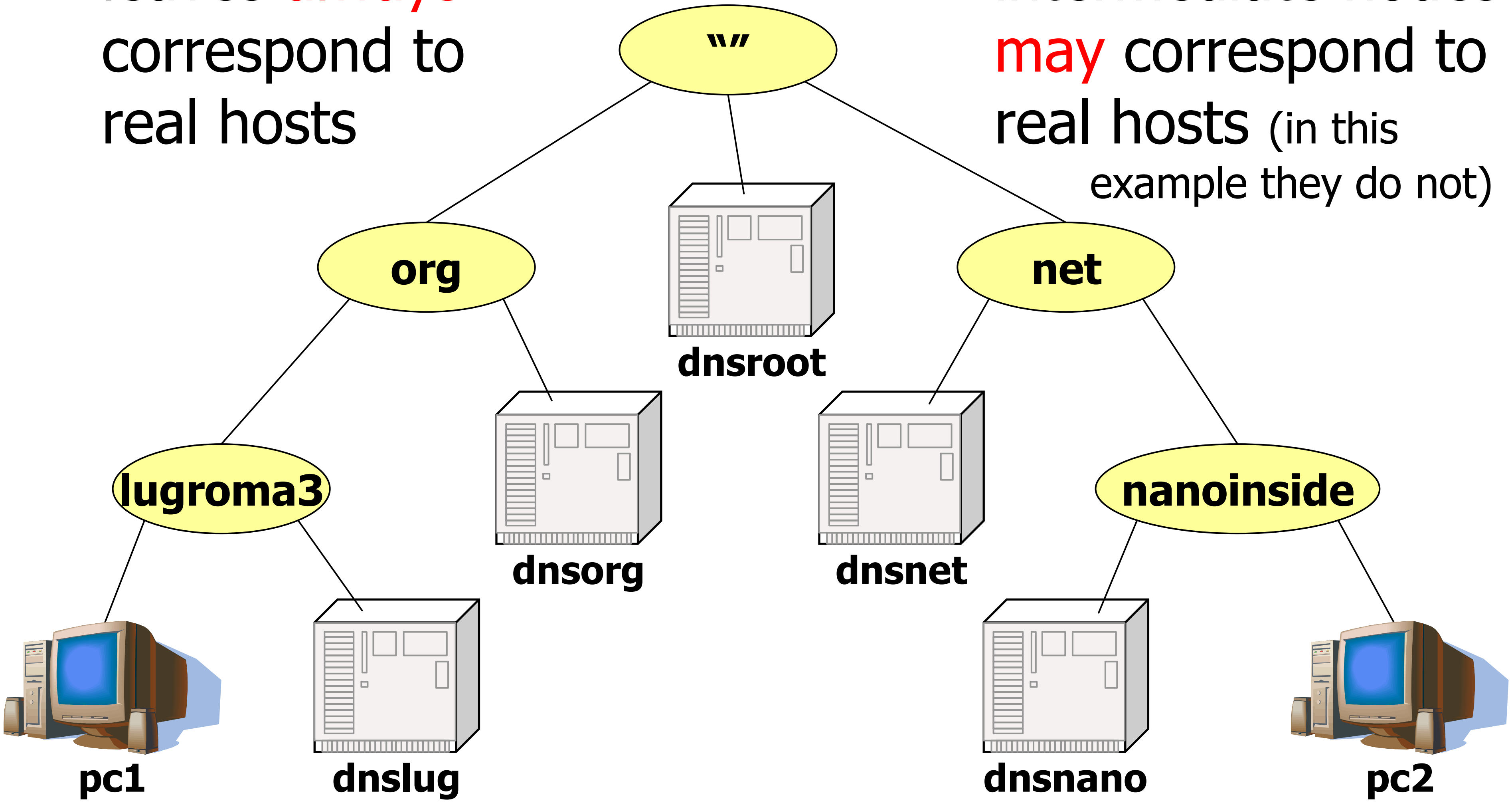
this node is **`nanoinside.net`**

this node is **`pc1.lugroma3.org`**

the dns name hierarchy

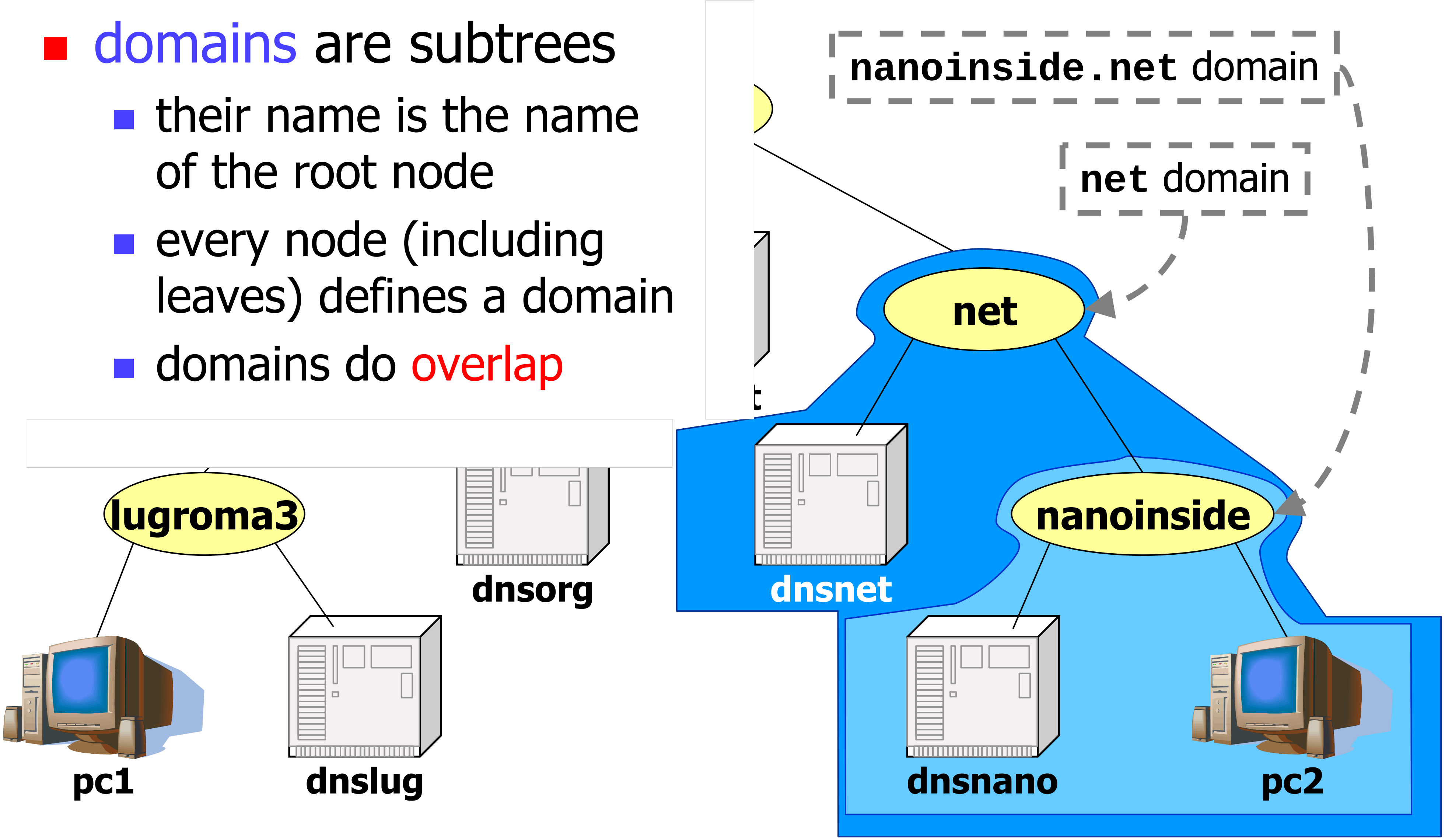
- leaves always correspond to real hosts

- intermediate nodes may correspond to real hosts (in this example they do not)



domains

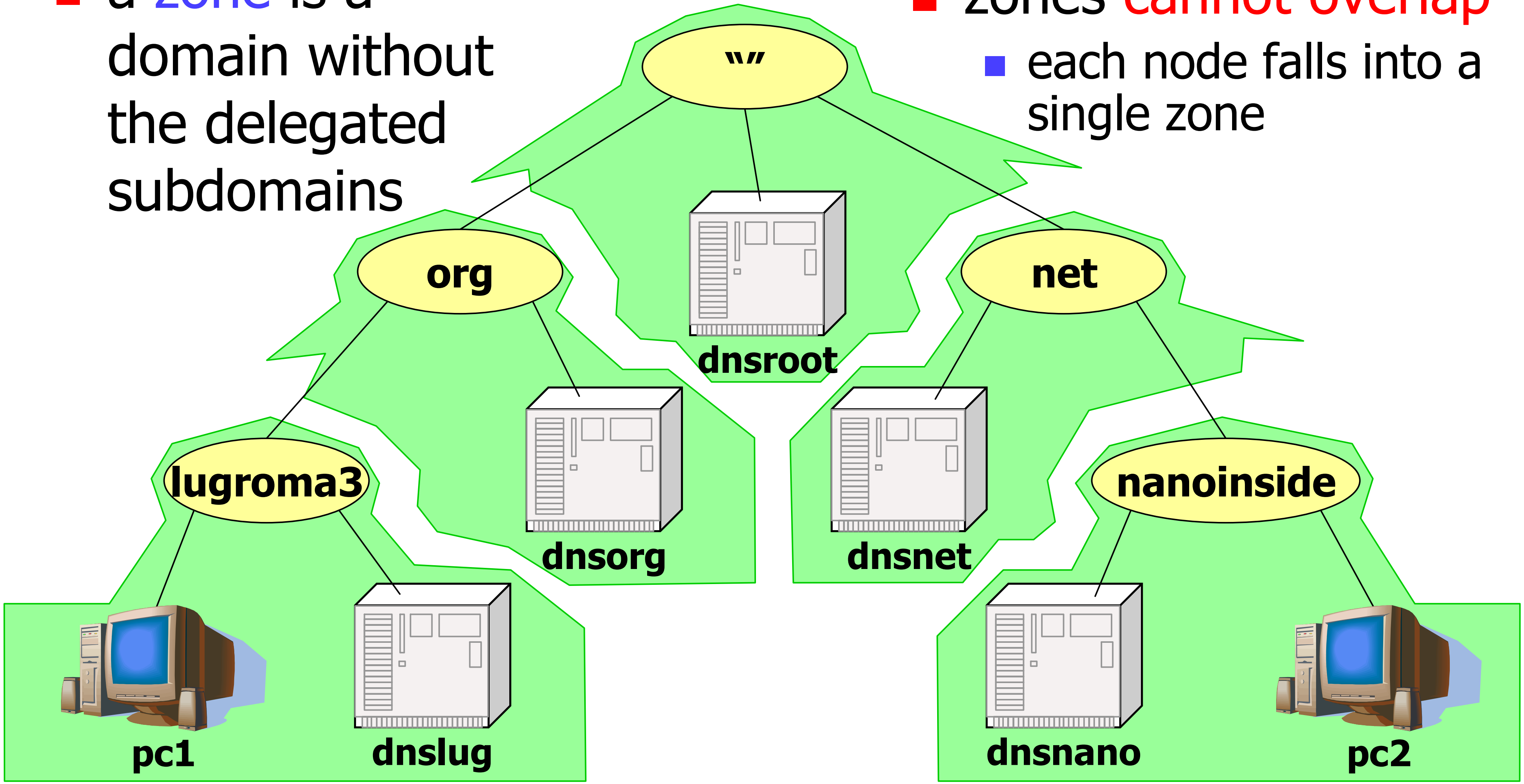
- domains are subtrees
  - their name is the name of the root node
  - every node (including leaves) defines a domain
  - domains do overlap



zones

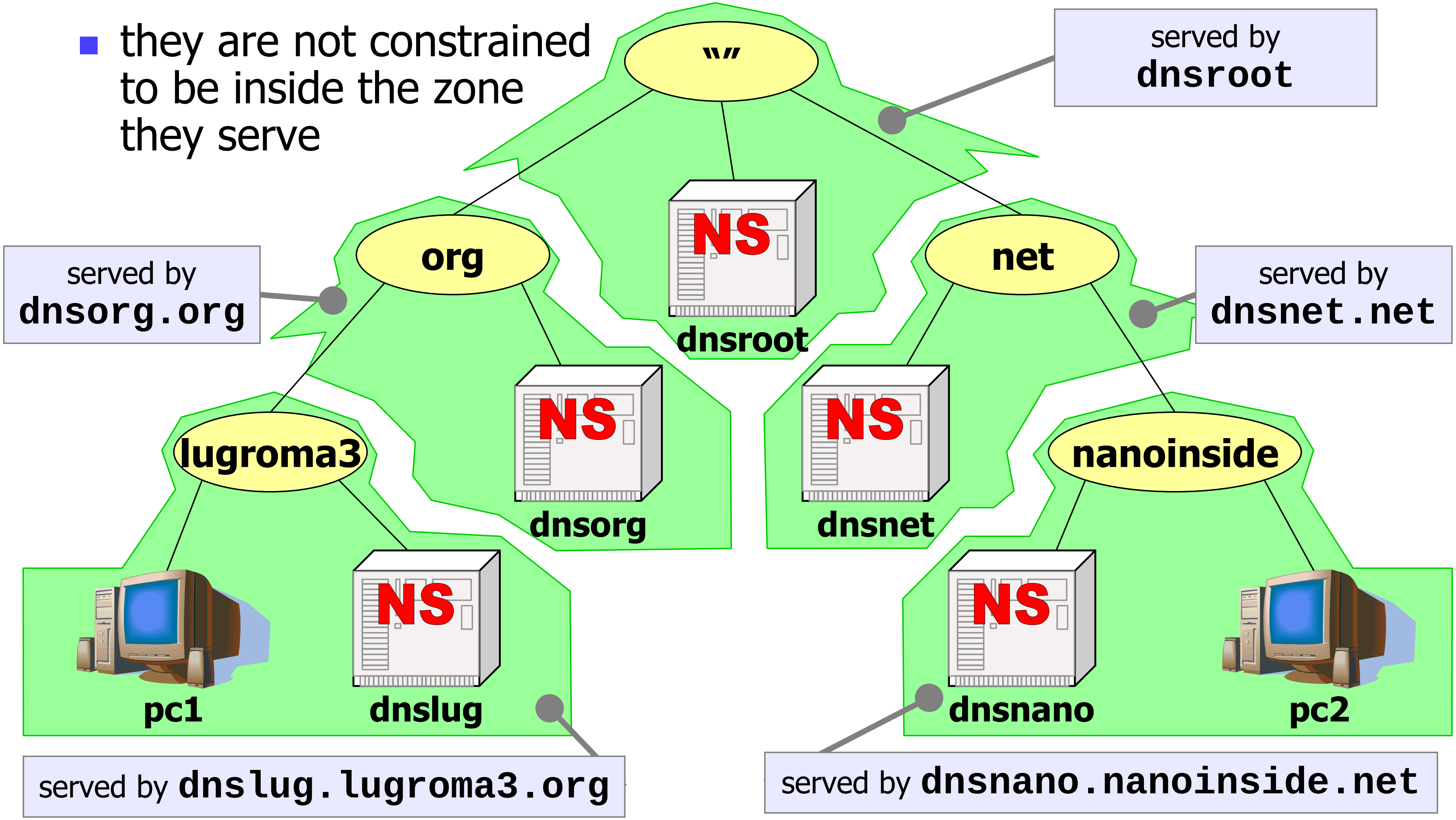
- a zone is a domain without the delegated subdomains

- zones cannot overlap
  - each node falls into a single zone



zones

- zones have name servers
  - they are not constrained to be inside the zone they serve



served by **`dnsroot`**

served by **`dnsorg.org`**

served by **`dnsnet.net`**

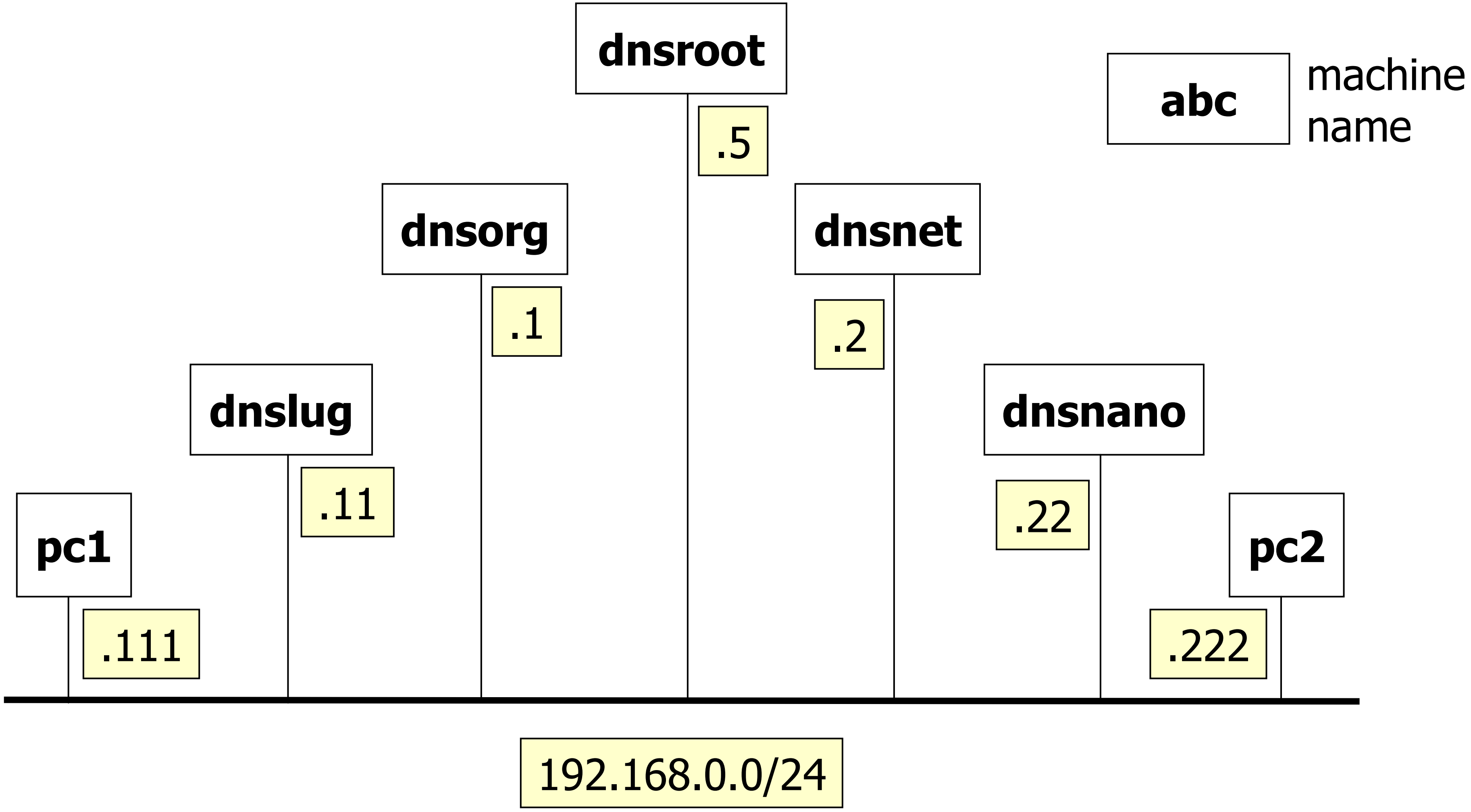
served by **`dnslug.lugroma3.org`**

served by **`dnsnano.nanoinside.net`**

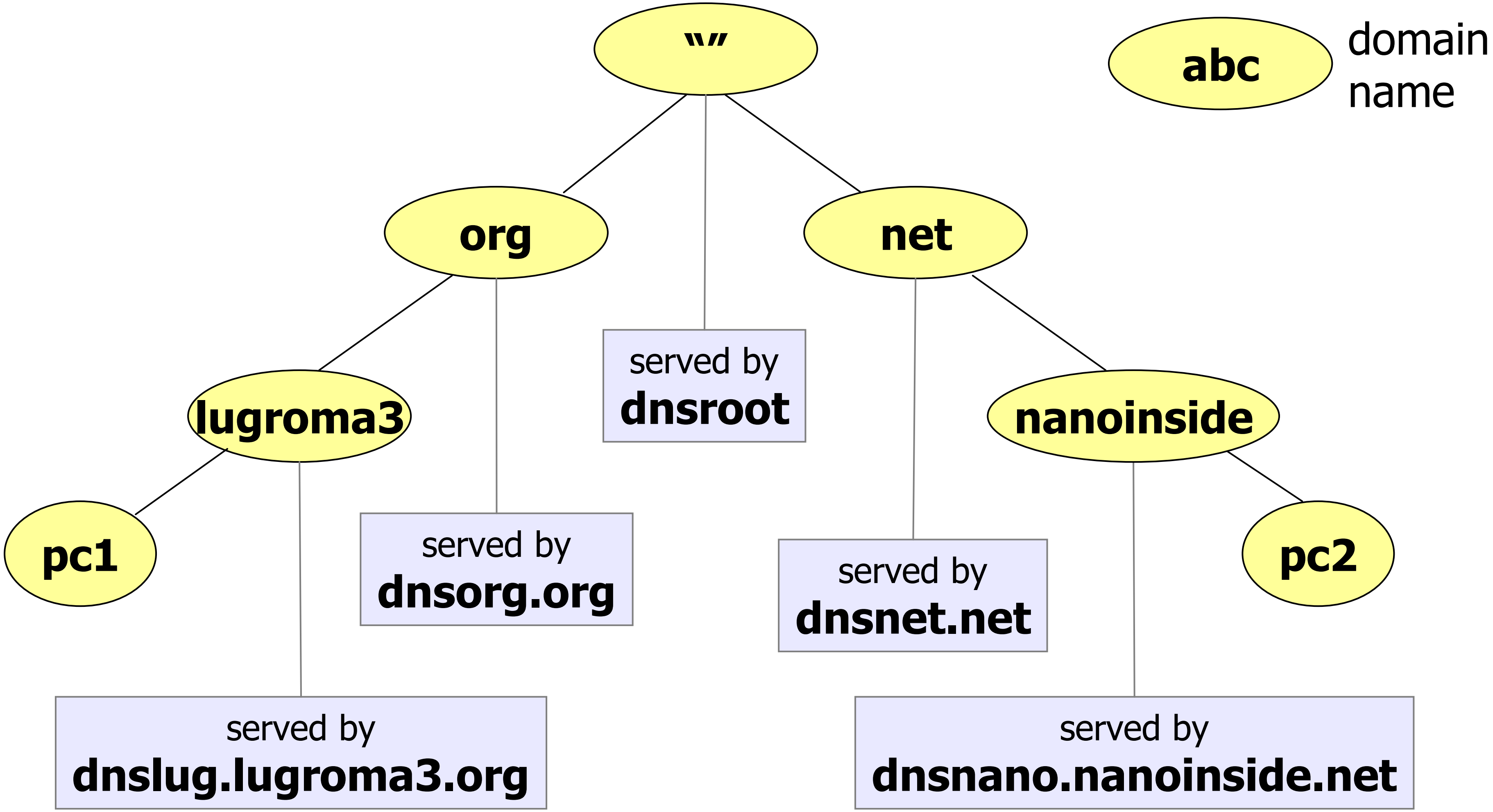
more about the dns

- the dns hierarchy is orthogonal with respect to the actual network topology
- in order to focus on the behavior of the dns we choose a flat topology, consisting of a single collision domain

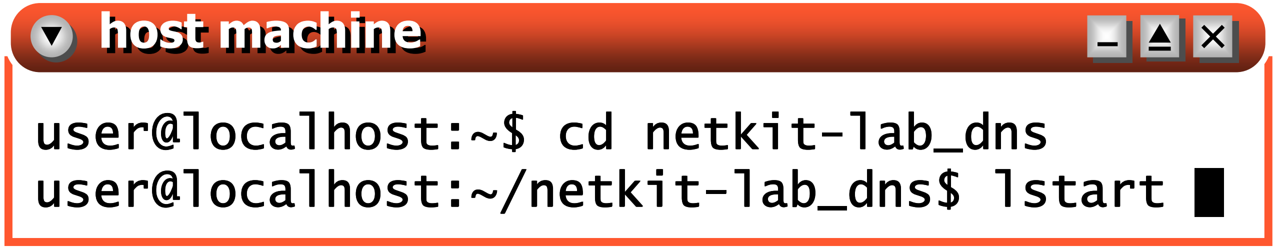
step 1 – network topology



step 1 – dns (zone) hierarchy



step 2 – starting the lab

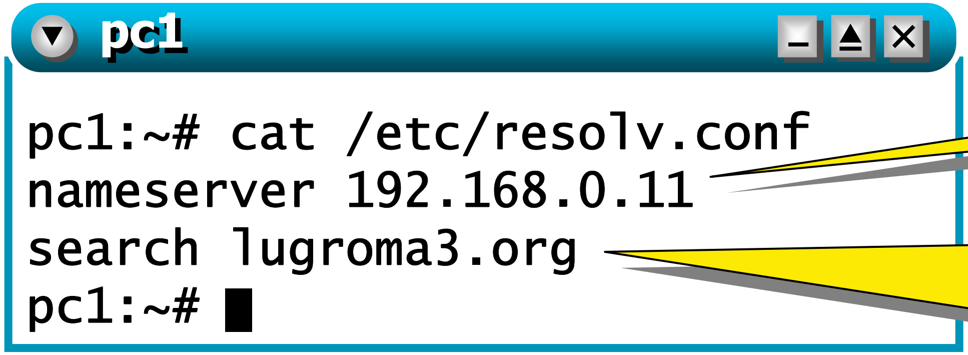


```
user@localhost:~$ cd netkit-lab_dns
user@localhost:~/netkit-lab_dns$ lstart █
```

- the lab is configured to
  - start all the 7 vms
  - automatically configure the network interfaces
  - automatically configure the name servers
  - automatically start the name server software (bind) on each name server

step 2 – exploring the configuration

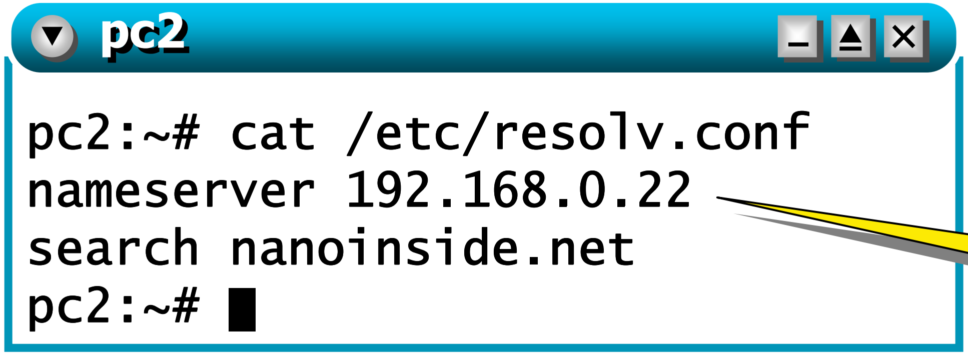
- configuration on the pcs consists of the specification of the default name server



```
pc1:~# cat /etc/resolv.conf
nameserver 192.168.0.11
search lugroma3.org
pc1:~# █
```

dnslug.lugroma3.org

suffix to append to unqualified names (e.g. asking to resolve **dummy** results in querying for **dummy.lugroma3.org**)



```
pc2:~# cat /etc/resolv.conf
nameserver 192.168.0.22
search nanoinside.net
pc2:~# █
```

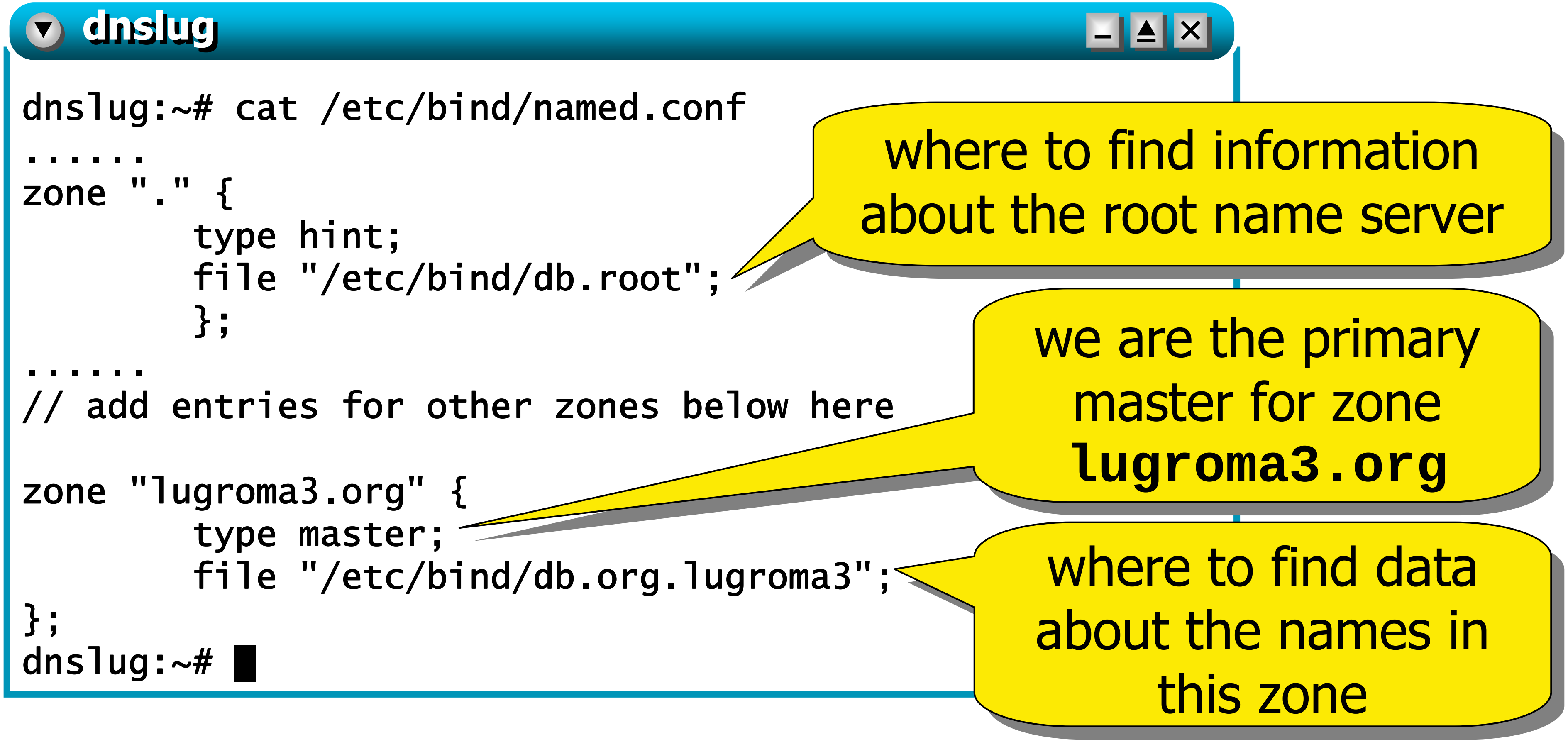
dnsnano.nanoinside.net

step 2 – exploring the configuration

- configuration on the name servers specifies
  - associations between zones and name servers
  - information about the root name servers
  - authoritative information
  - associations between names and ip addresses

step 2 – exploring the configuration

- configuration on the name servers specifies
  - associations between zones and name servers



```
dnslug:~# cat /etc/bind/named.conf
......
zone "." {
        type hint;
        file "/etc/bind/db.root";
        };
......
// add entries for other zones below here

zone "lugroma3.org" {
        type master;
        file "/etc/bind/db.org.lugroma3";
};
dnslug:~# █
```

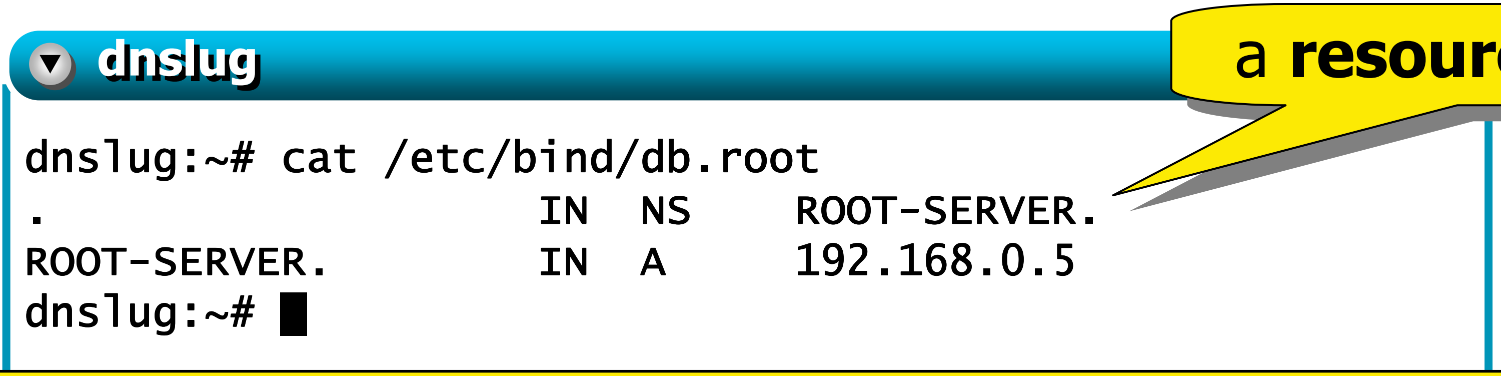
where to find information about the root name server

we are the primary master for zone **`lugroma3.org`**

where to find data about the names in this zone

step 2 – exploring the configuration

- configuration on the name servers specifies
  - information about the root name servers



```
dnslug:~# cat /etc/bind/db.root
.                       IN  NS    ROOT-SERVER.
ROOT-SERVER.            IN  A     192.168.0.5
dnslug:~# █
```

a **resource record**

format of a resource record

`<domain> <class> <type> <rdata>`

`domain`: the record owner (=domain to which the record refers)

`class`: usually IN (=Internet system); may be HS (=hesiod) or CH (=chaos)

`type`: see next slide...

`rdata`: record data (depends on the record type)

step 2 – exploring the configuration

available record types

A **a host address.**

A6 **an IPv6 address.**

AAAA **Obsolete format of IPv6 address**

AFSDB (x) location of AFS database servers. Experimental.

CERT holds a digital certificate.

CNAME identifies the canonical name of an alias.

DNAME for delegation of reverse addresses. Replaces the domain name specified with another name to be looked up. Described in RFC 2672.

GPOS Specifies the global position. Superseded by LOC.

HINFO identifies the CPU and OS used by a host.

ISDN (x) representation of ISDN addresses. Experimental.

KEY stores a public key associated with a DNS name.

KX identifies a key exchanger for this DNS name.

LOC (x) for storing GPS info. See RFC 1876. Experimental.

MX **identifies a mail exchange for the domain. See RFC 974 for details.**

NAPTR name authority pointer.

NSAP a network service access point.

NS **the authoritative nameserver for the domain.**

NXT used in DNSSEC to securely indicate that RRs with an owner name in a certain name interval do not exist in a zone and indicate what R

PTR **a pointer to another part of the domain name space.**

PX provides mappings between RFC 822 and X.400 addresses.

RP (x) information on persons responsible for the domain. Experimental.

RT (x) route-through binding for hosts that do not have their own direct wide area network addresses. Experimental.

SIG ("signature") contains data authenticated in the secure DNS. See RFC 2535 for details.

SOA **identifies the start of a zone of authority.**

SRV information about well known network services (replaces WKS).

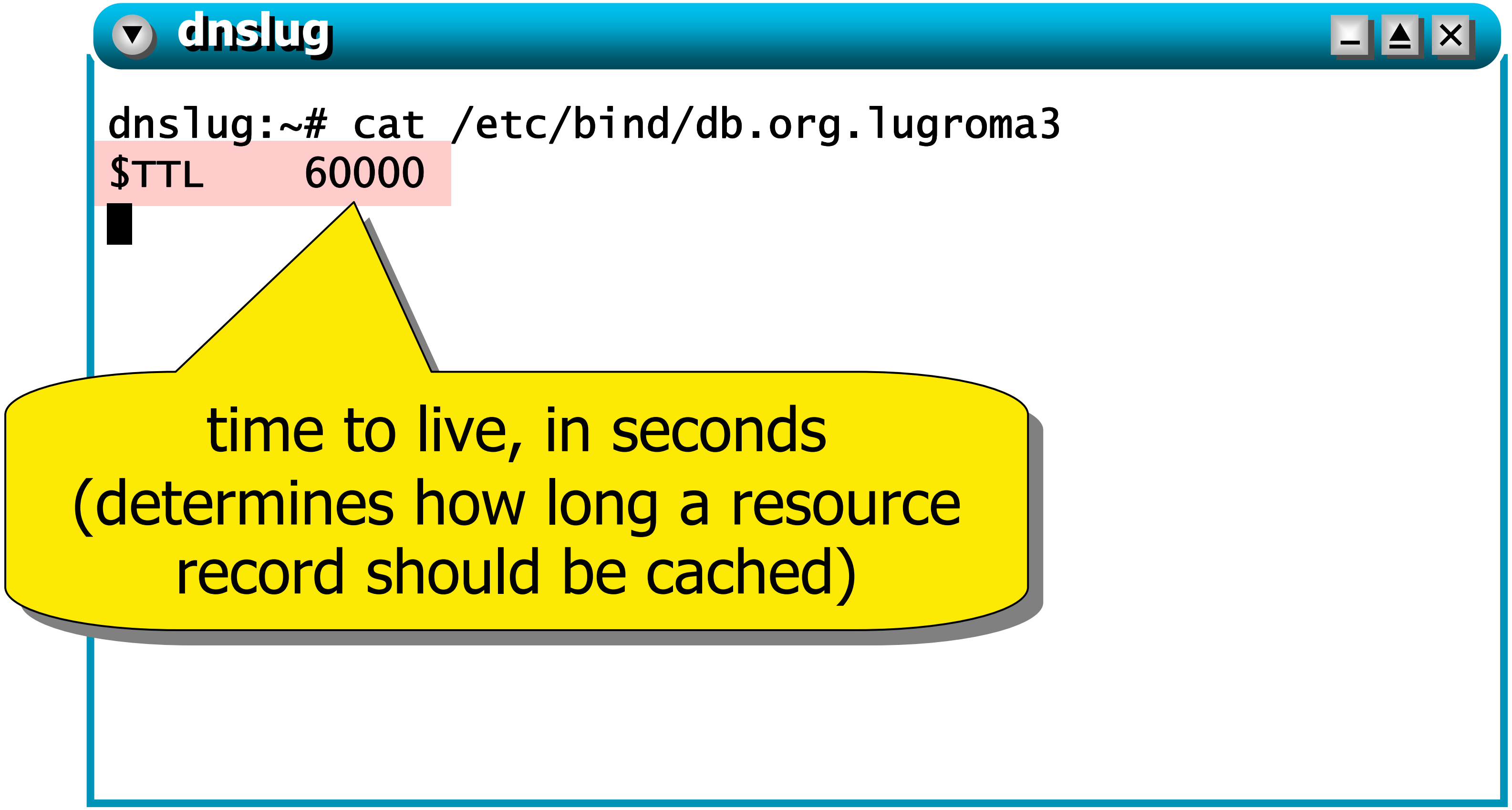
TXT text records.

WKS (h) information about which well known network services, such as SMTP, that a domain supports. Historical, replaced by newer RR SRV.

X25 (x) representation of X.25 network addresses. Experimental

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information

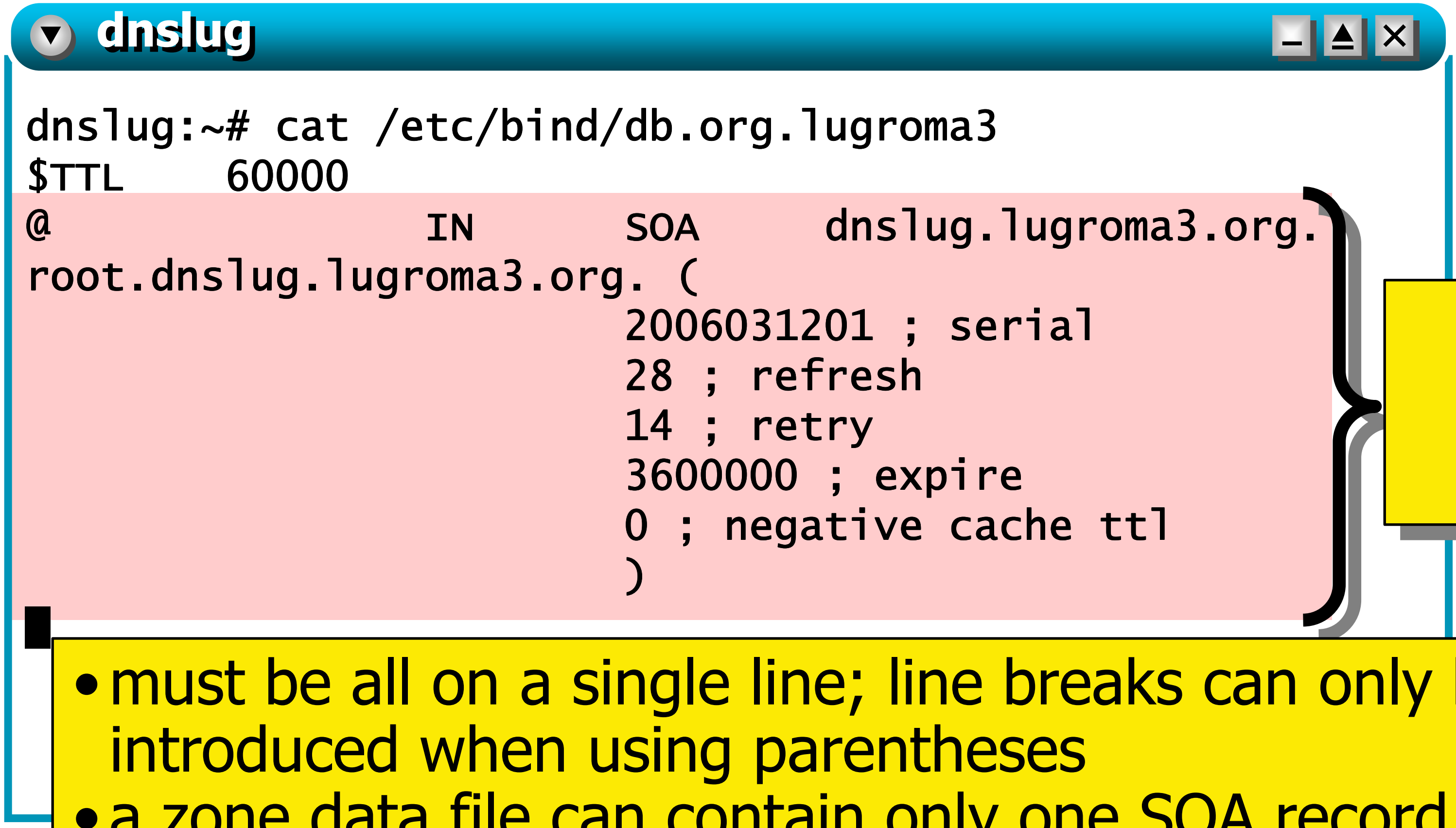


```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
```

time to live, in seconds (determines how long a resource record should be cached)

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



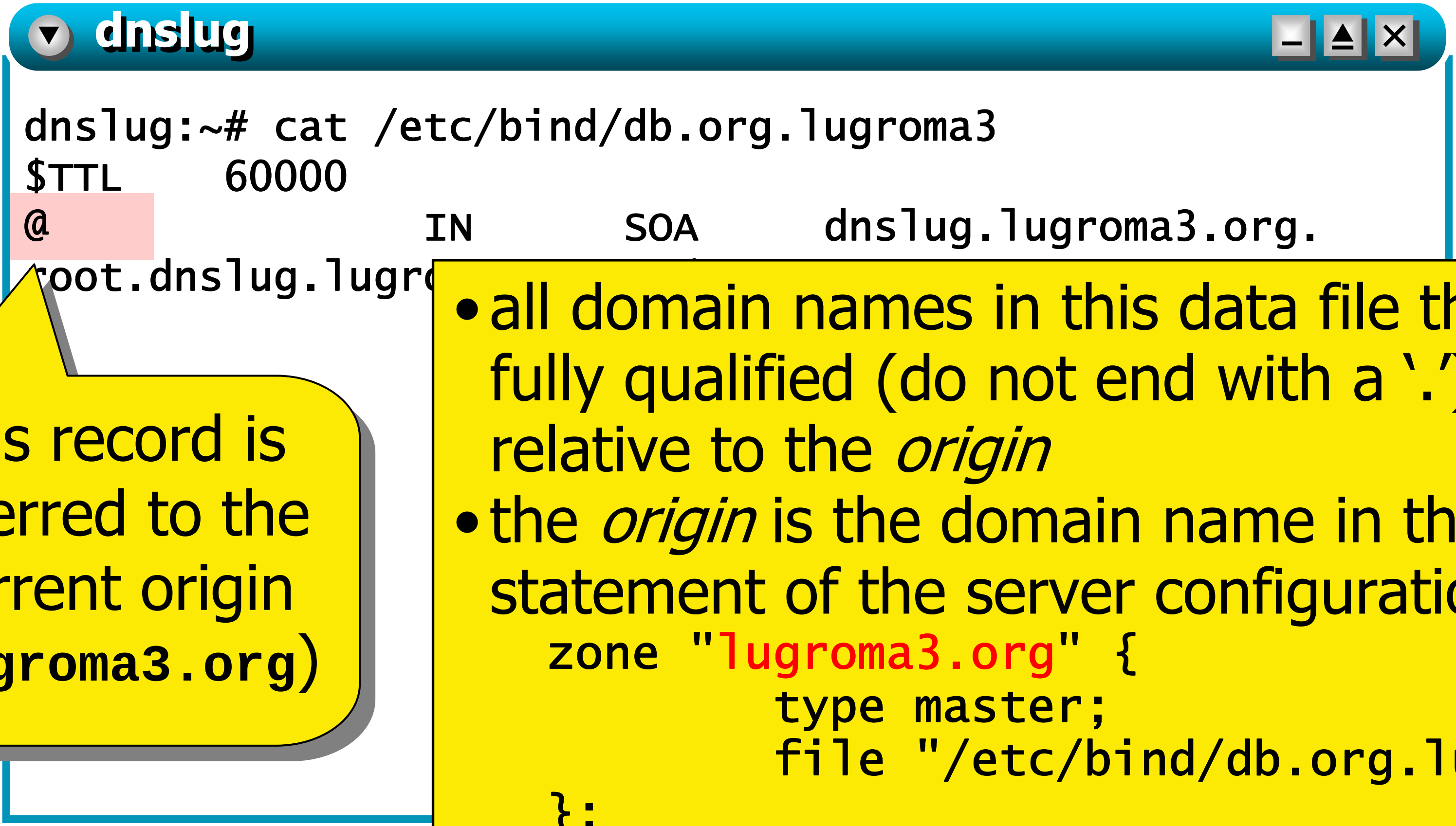
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                        28 ; refresh
                        14 ; retry
                        3600000 ; expire
                        0 ; negative cache ttl
                        )
```

Start of Authority record

- must be all on a single line; line breaks can only be introduced when using parentheses
- a zone data file can contain only one SOA record

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugro
```

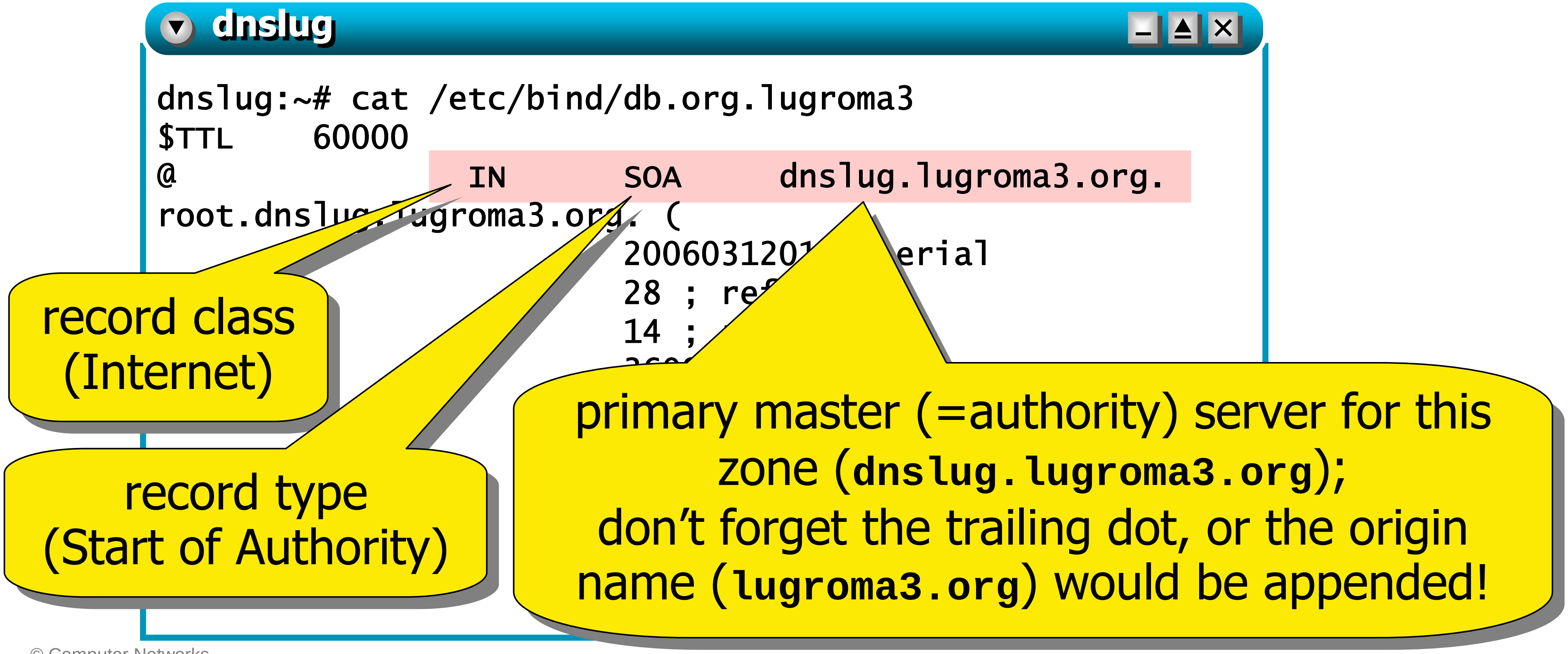
this record is referred to the current origin (`lugroma3.org`)

- all domain names in this data file that are not fully qualified (do not end with a '.') are relative to the *origin*
- the *origin* is the domain name in the *zone* statement of the server configuration file:

```
zone "lugroma3.org" {
        type master;
        file "/etc/bind/db.org.lugroma3";
};
```

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN        SOA       dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                          2006031201       erial
                          28 ; ref
                          14 ;
                          360
```

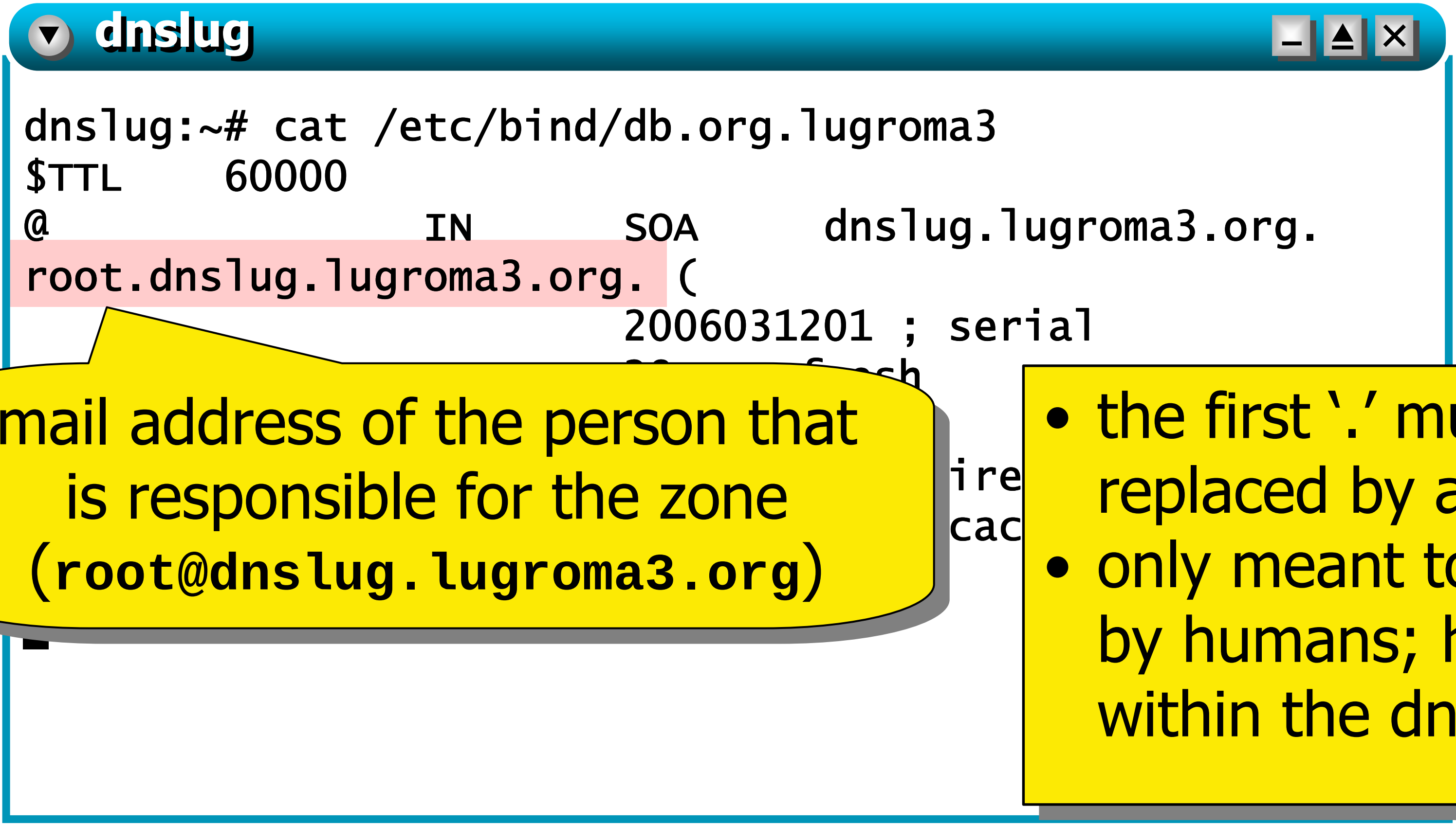
record class (Internet)

record type (Start of Authority)

primary master (=authority) server for this zone (**`dnslug.lugroma3.org`**); don't forget the trailing dot, or the origin name (**`lugroma3.org`**) would be appended!

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



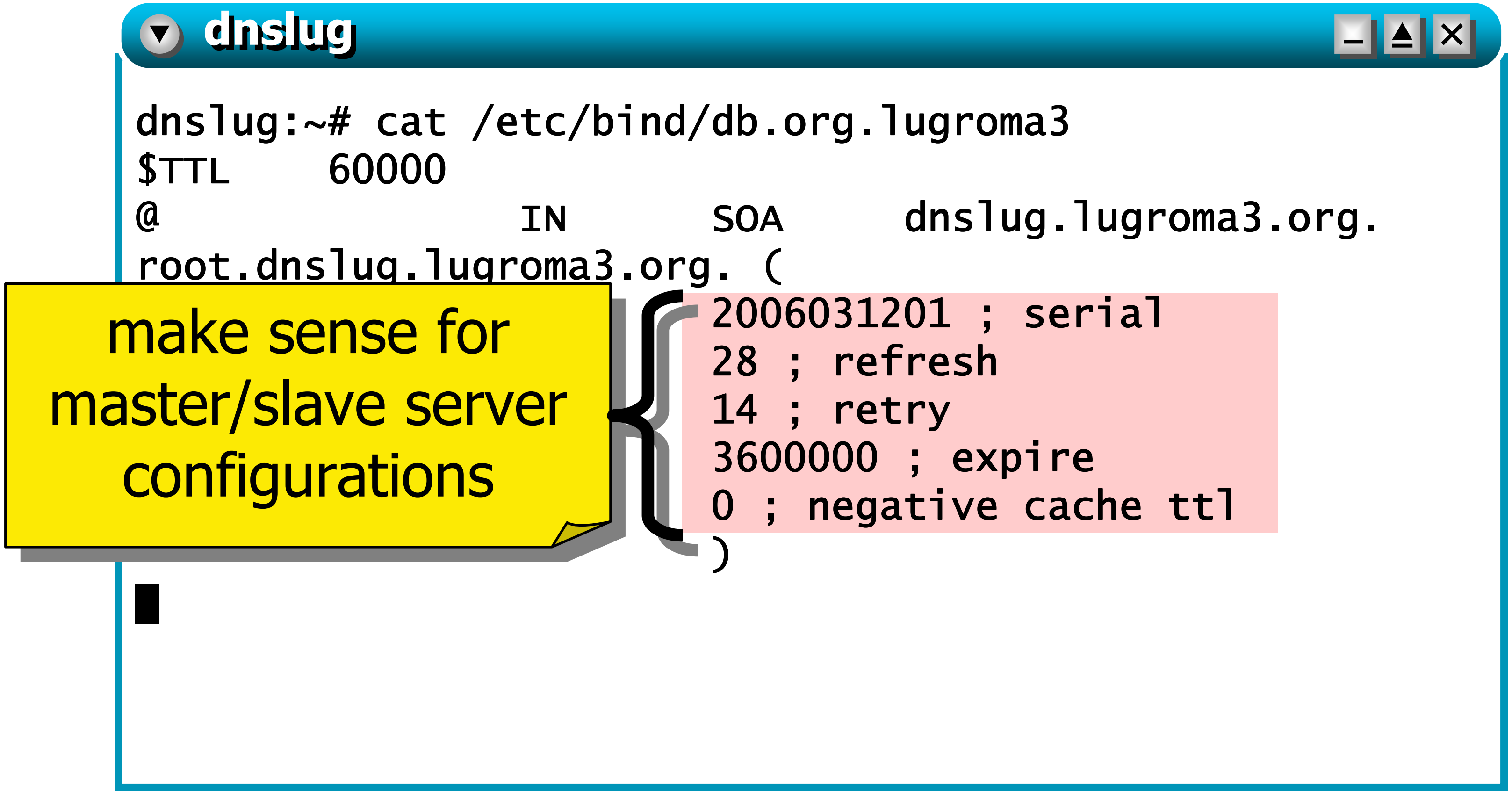
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                                      ire
                                      cac
```

mail address of the person that is responsible for the zone (**root@dnslug.lugroma3.org**)

- the first '.' must be replaced by a '@'
- only meant to be used by humans; has no use within the dns service

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information

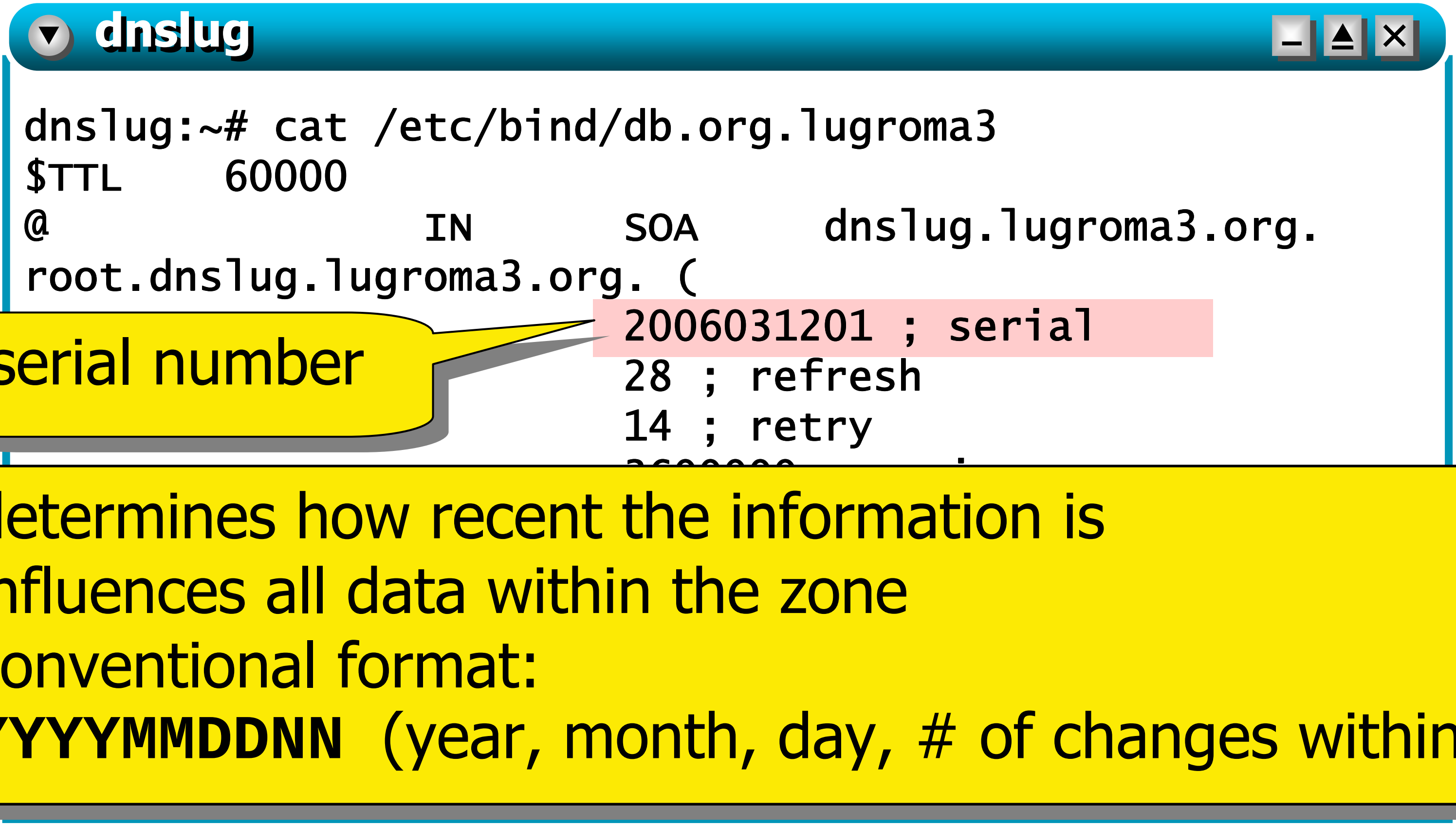


```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                                2006031201 ; serial
                                28 ; refresh
                                14 ; retry
                                3600000 ; expire
                                0 ; negative cache ttl
                                )
█
```

make sense for master/slave server configurations

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



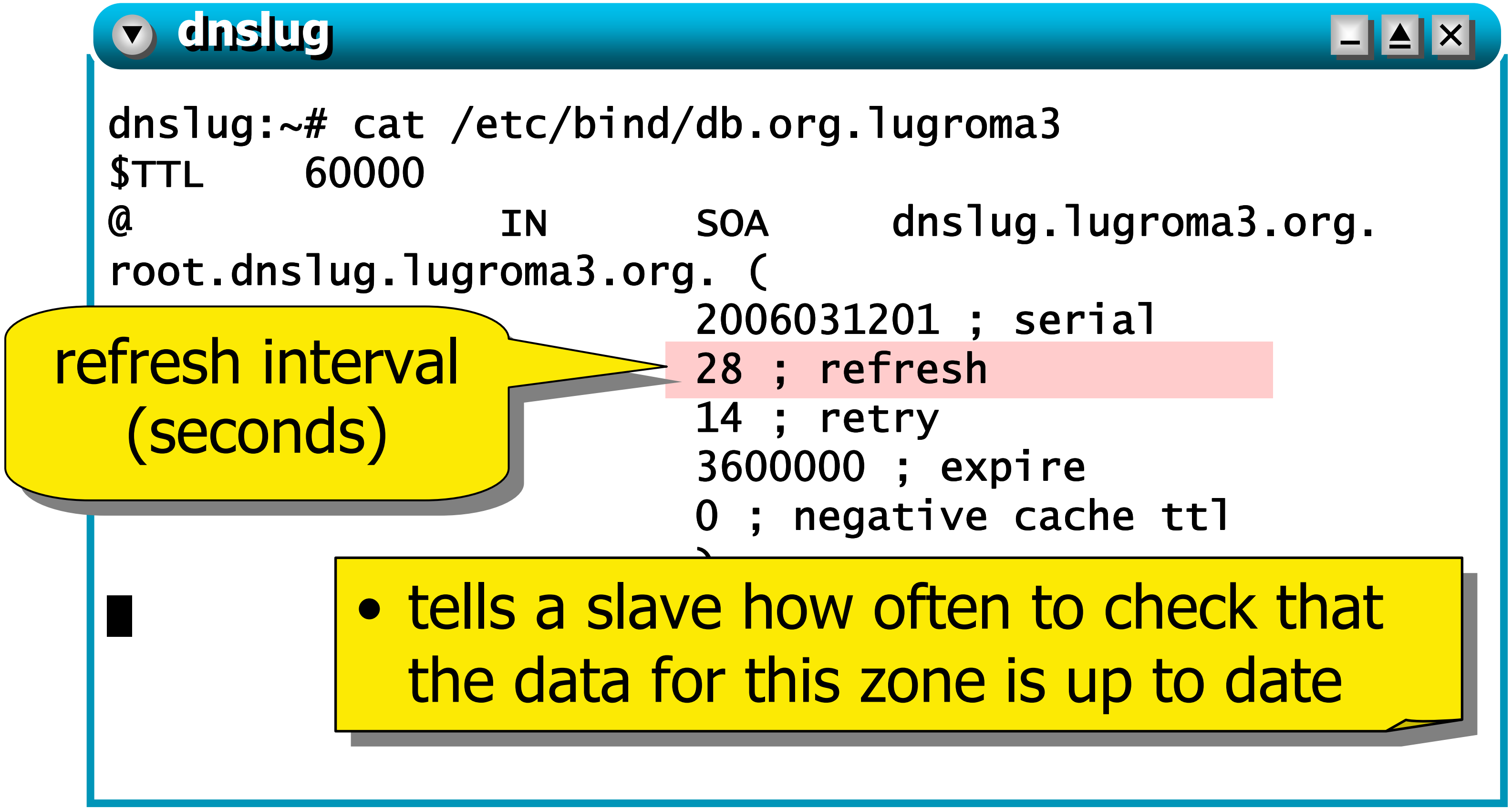
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@                IN       SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                                  2006031201 ; serial
                                  28 ; refresh
                                  14 ; retry
```

serial number

- determines how recent the information is
- influences all data within the zone
- conventional format:
  YYYYMMDDNN (year, month, day, # of changes within that day)

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



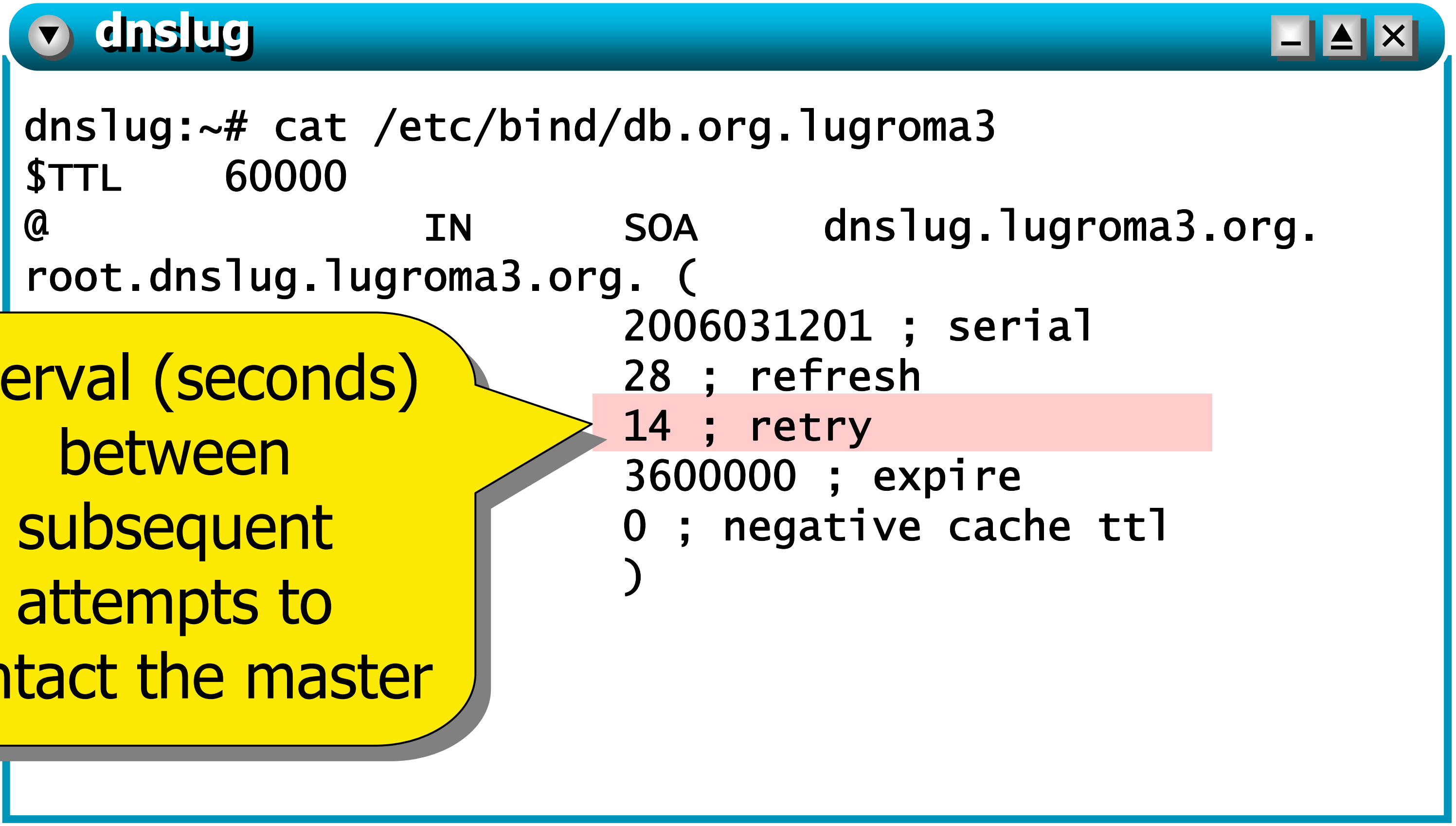
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                        28 ; refresh
                        14 ; retry
                        3600000 ; expire
                        0 ; negative cache ttl
                        )
█
```

refresh interval (seconds)

- tells a slave how often to check that the data for this zone is up to date

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information

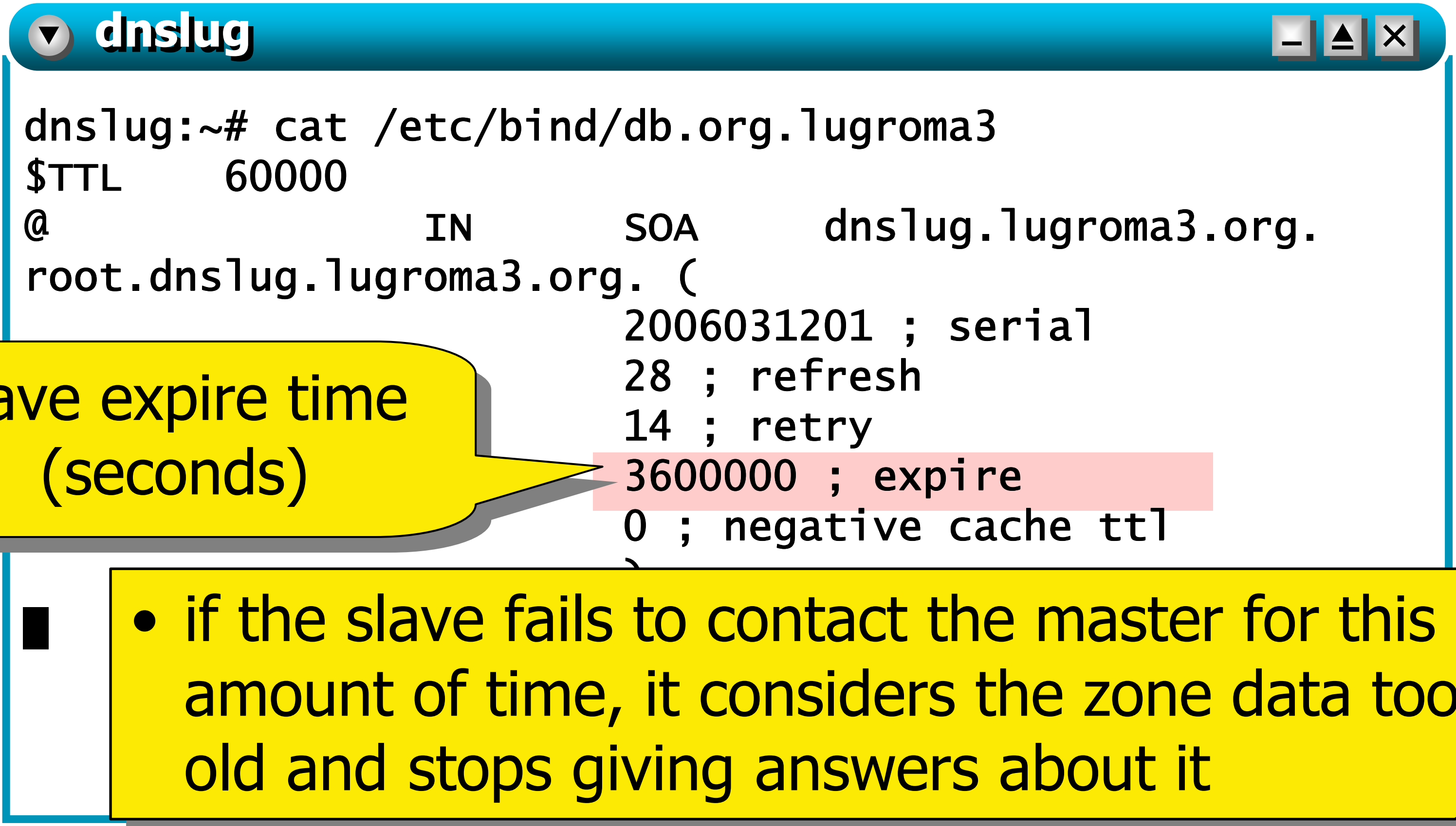


```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                        28 ; refresh
                        14 ; retry
                        3600000 ; expire
                        0 ; negative cache ttl
                        )
```

interval (seconds) between subsequent attempts to contact the master

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information



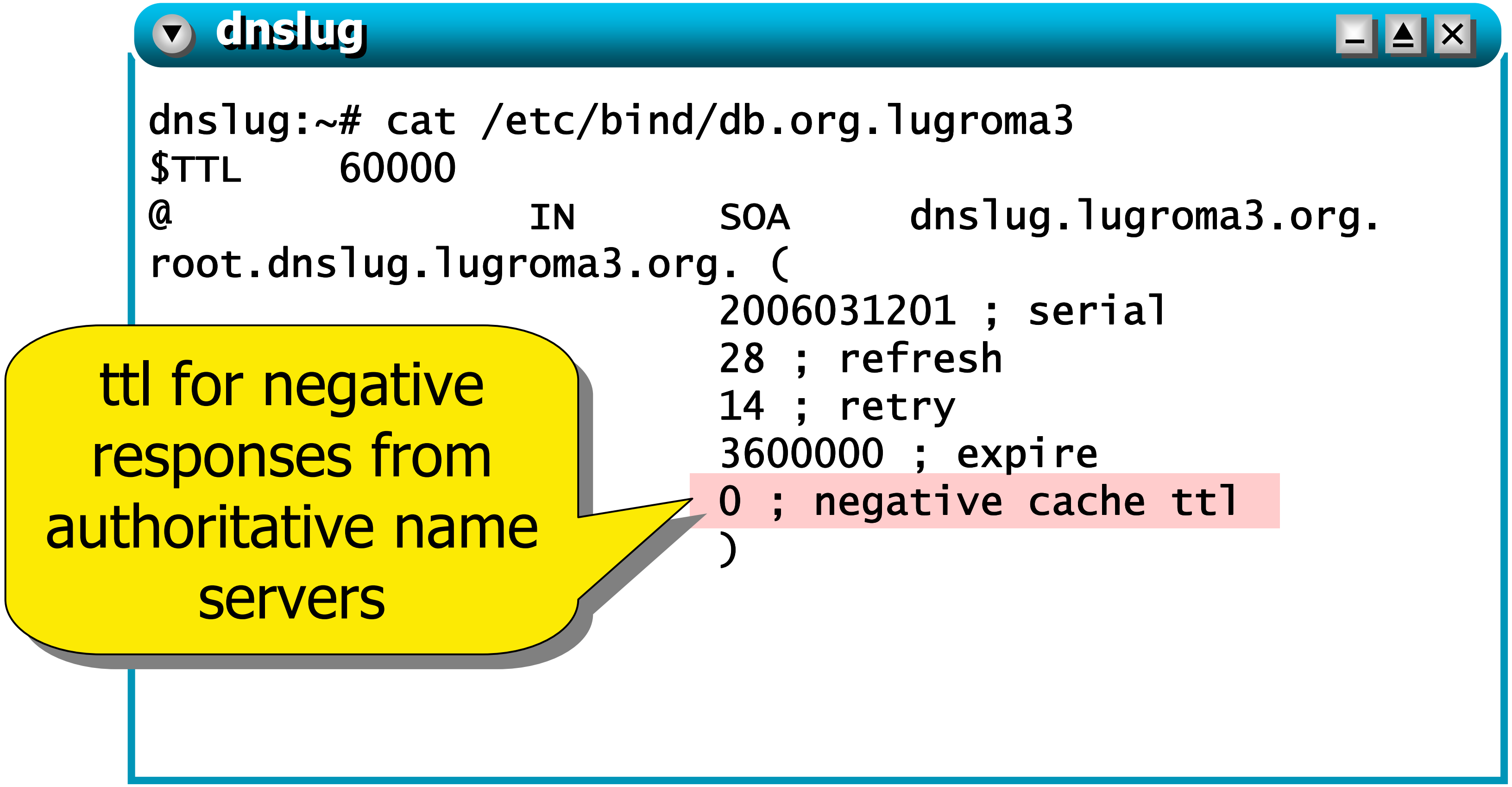
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                        28 ; refresh
                        14 ; retry
                        3600000 ; expire
                        0 ; negative cache ttl
                        )
█
```

slave expire time (seconds)

- if the slave fails to contact the master for this amount of time, it considers the zone data too old and stops giving answers about it

step 2 – exploring the configuration

- configuration on the name servers specifies
  - authoritative information

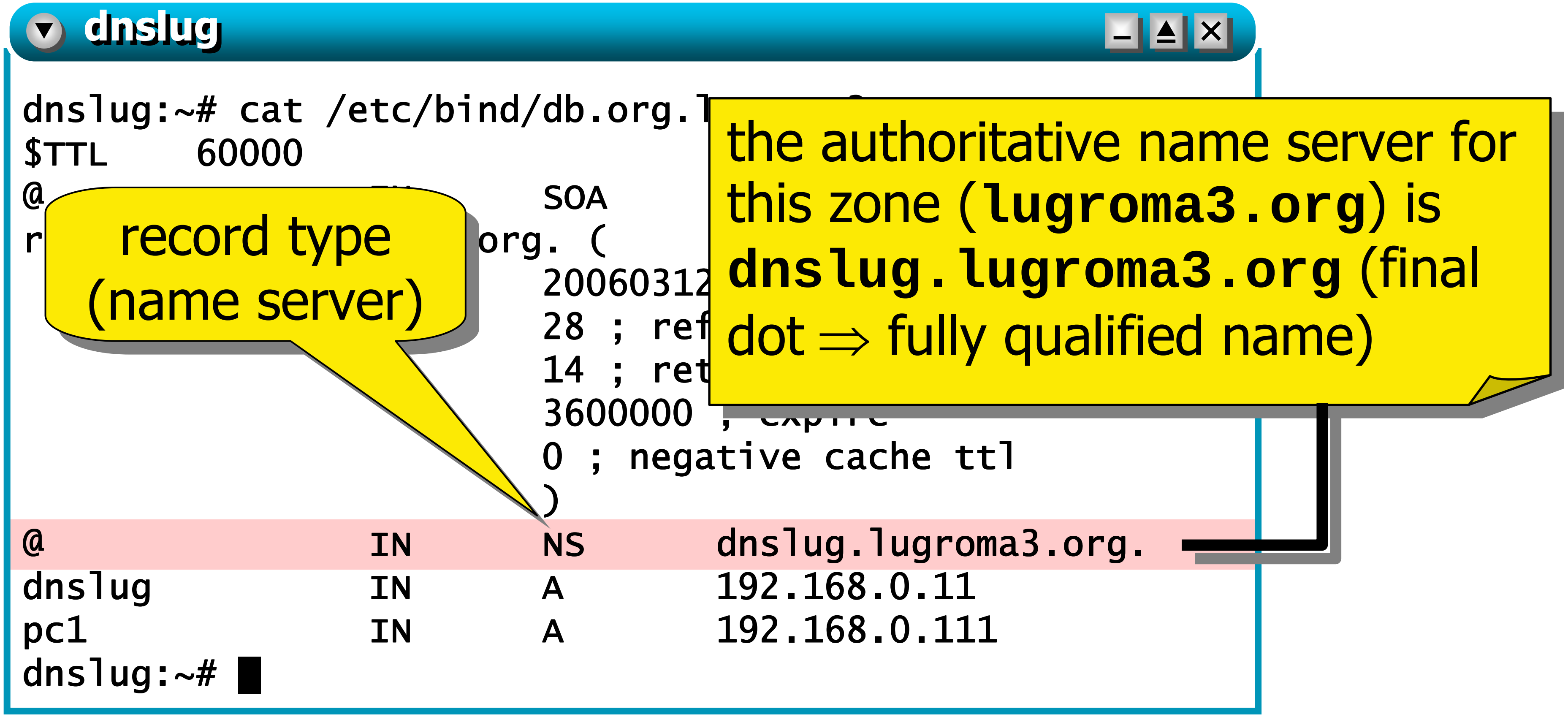


```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug.lugroma3.org.
root.dnslug.lugroma3.org. (
                        2006031201 ; serial
                        28 ; refresh
                        14 ; retry
                        3600000 ; expire
                        0 ; negative cache ttl
                        )
```

ttl for negative responses from authoritative name servers

step 2 – exploring the configuration

- configuration on the name servers specifies
  - associations between names and ip addresses



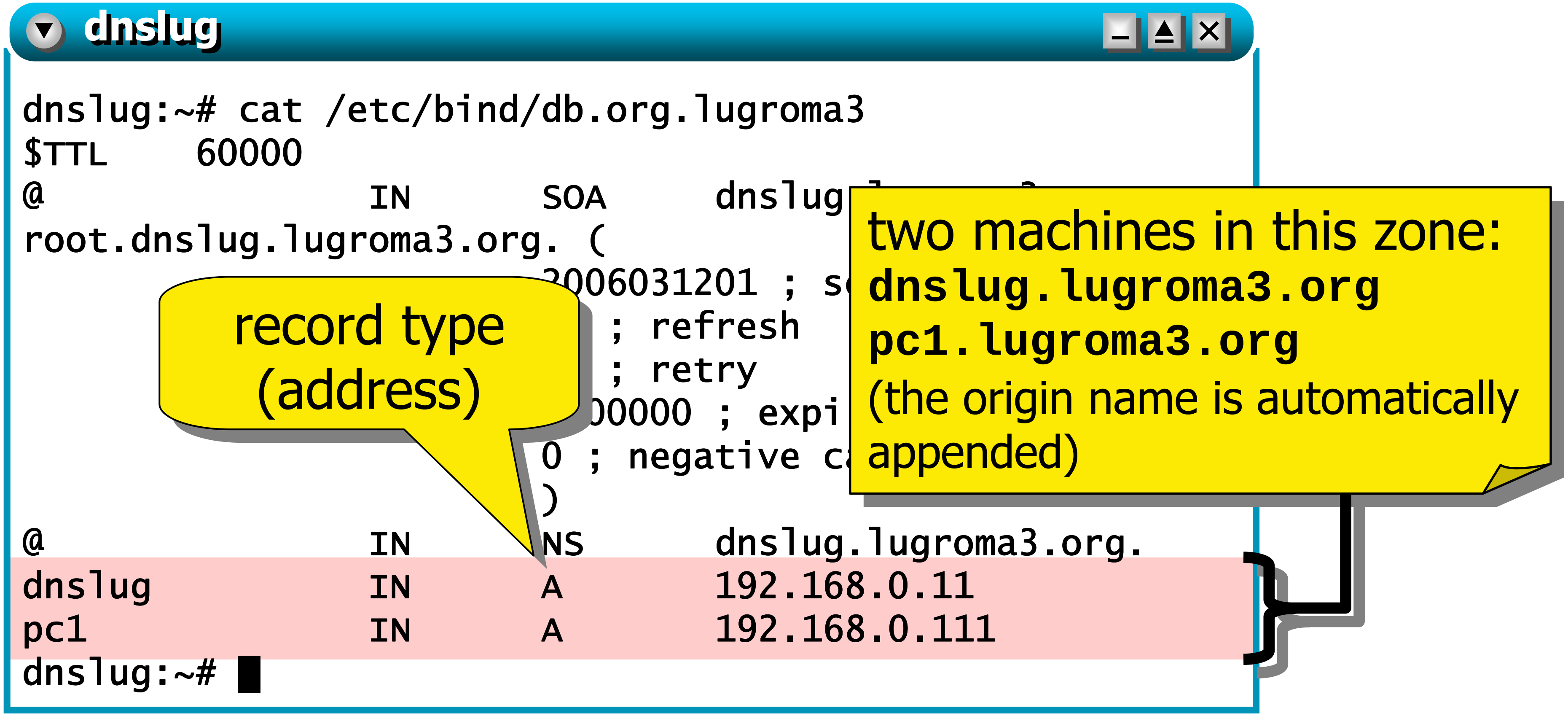
```
dnslug:~# cat /etc/bind/db.org.l
$TTL     60000
@                          SOA
r                     org. (
                           20060312
                           28 ; ref
                           14 ; ret
                           3600000 ; expire
                           0 ; negative cache ttl
                           )
@                IN        NS        dnslug.lugroma3.org.
dnslug           IN        A         192.168.0.11
pc1              IN        A         192.168.0.111
dnslug:~# █
```

record type (name server)

the authoritative name server for this zone (**lugroma3.org**) is **dnslug.lugroma3.org** (final dot ⇒ fully qualified name)

step 2 – exploring the configuration

- configuration on the name servers specifies
  - associations between names and ip addresses



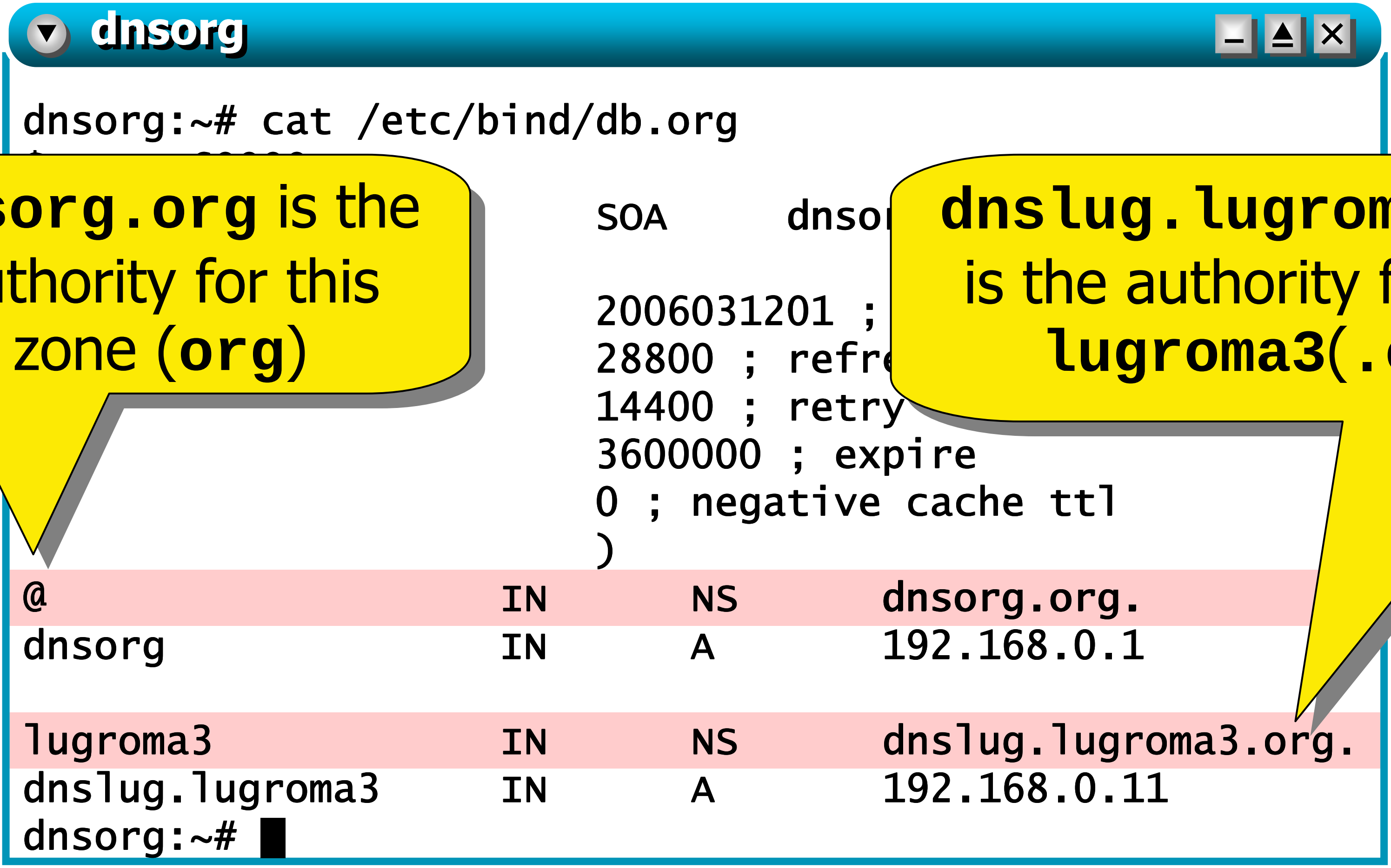
```
dnslug:~# cat /etc/bind/db.org.lugroma3
$TTL    60000
@               IN      SOA     dnslug
root.dnslug.lugroma3.org. (
                        2006031201 ; s
                           ; refresh
                           ; retry
                          00000 ; expi
                        0 ; negative c
                        )
@               IN      NS      dnslug.lugroma3.org.
dnslug          IN      A       192.168.0.11
pc1             IN      A       192.168.0.111
dnslug:~# █
```

record type (address)

two machines in this zone:
`dnslug.lugroma3.org`
`pc1.lugroma3.org`
(the origin name is automatically appended)

step 2 – exploring the configuration

- configuration on the name servers may specify
  - an authority for a subdomain

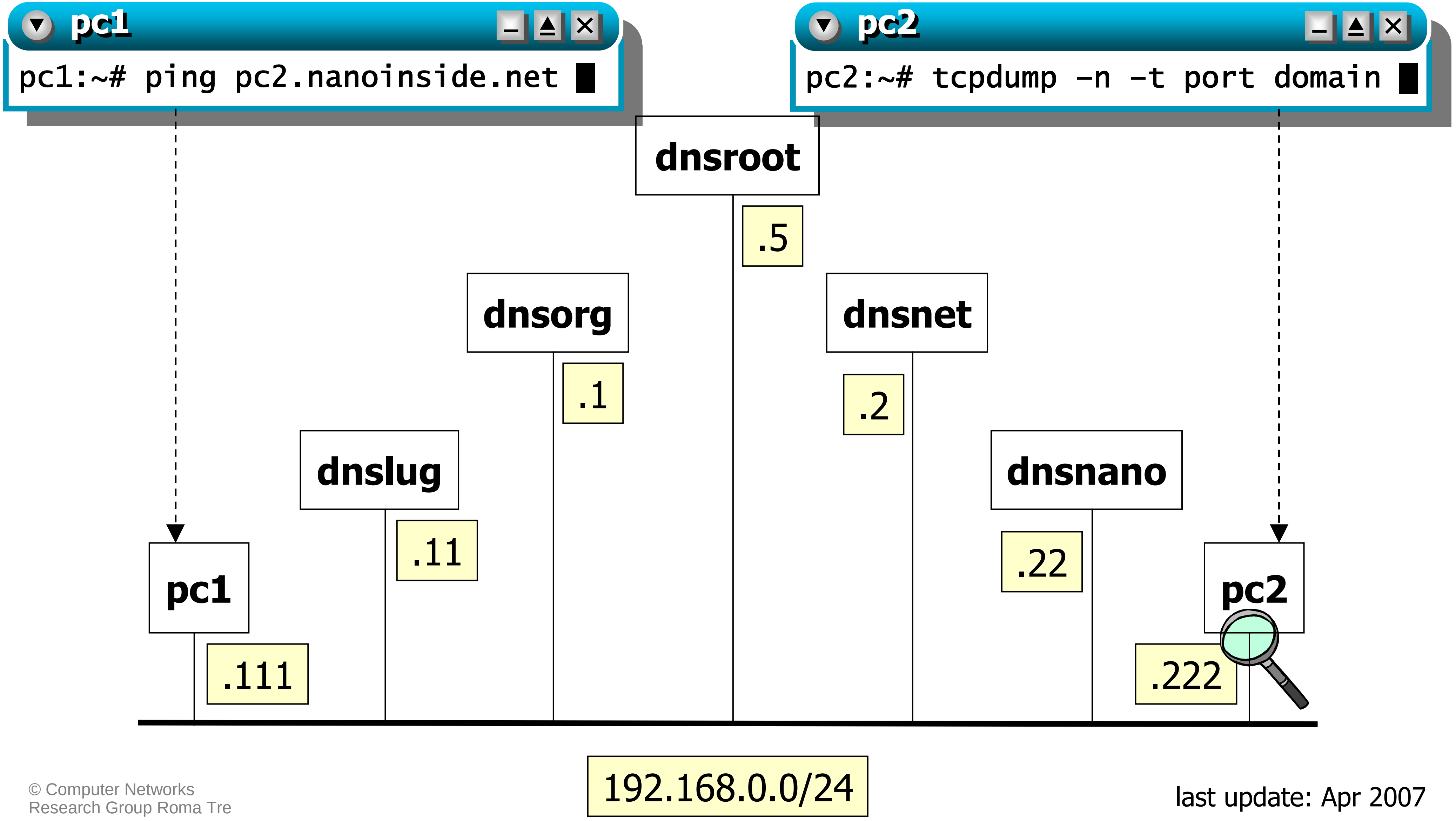


```
dnsorg:~# cat /etc/bind/db.org

                         SOA     dnso

                         2006031201 ;
                         28800 ; refre
                         14400 ; retry
                         3600000 ; expire
                         0 ; negative cache ttl
                         )
@                    IN       NS      dnsorg.org.
dnsorg               IN       A       192.168.0.1

lugroma3             IN       NS      dnslug.lugroma3.org.
dnslug.lugroma3      IN       A       192.168.0.11
dnsorg:~# █
```

dnsorg.org is the authority for this zone (**org**)

dnslug.lugroma3.org is the authority for zone **lugroma3**(**.org**)

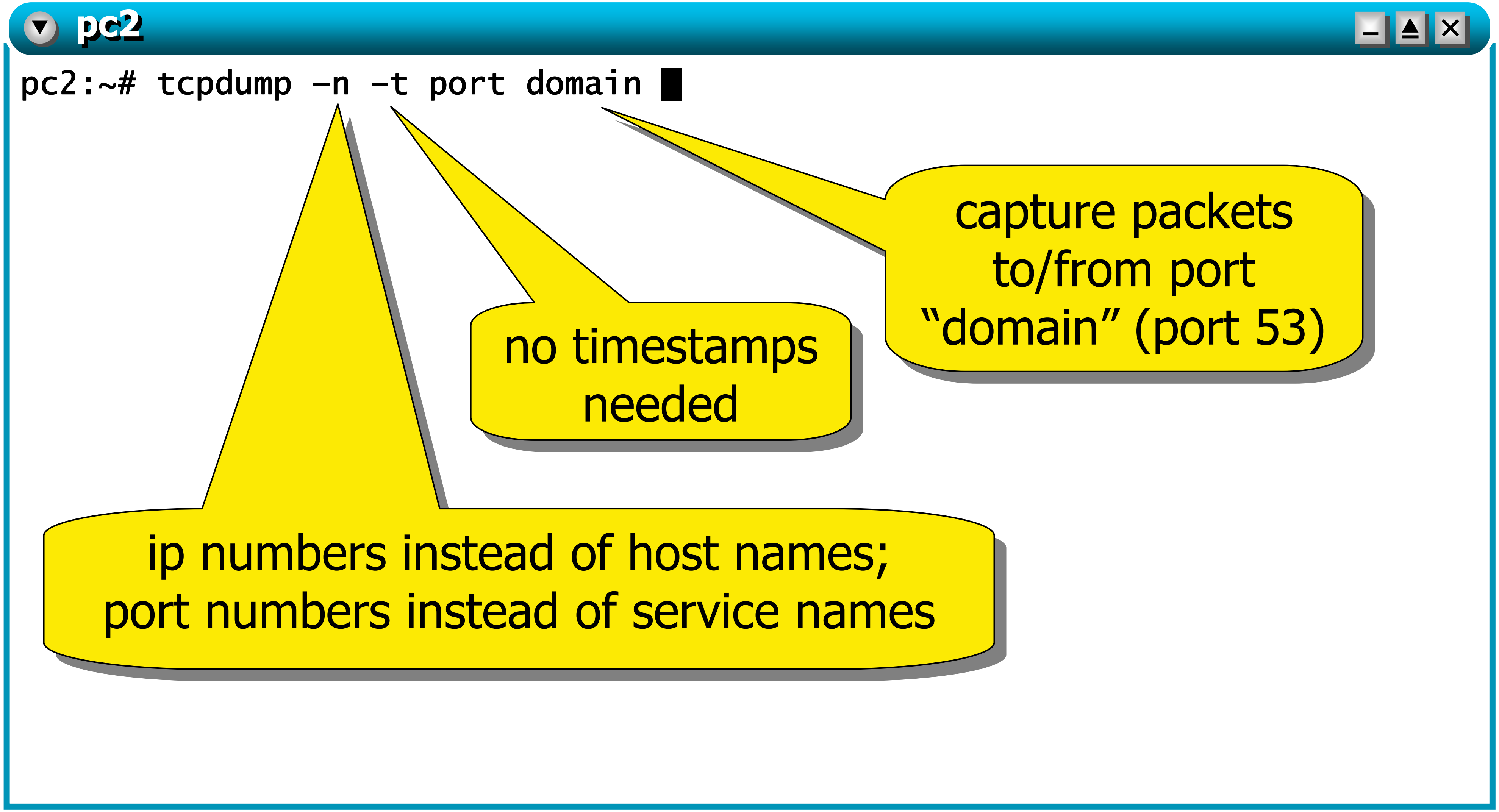
step 3 – experiment setting



```
pc1:~# ping pc2.nanoinside.net
```

```
pc2:~# tcpdump -n -t port domain
```

step 3 – the sniffer output



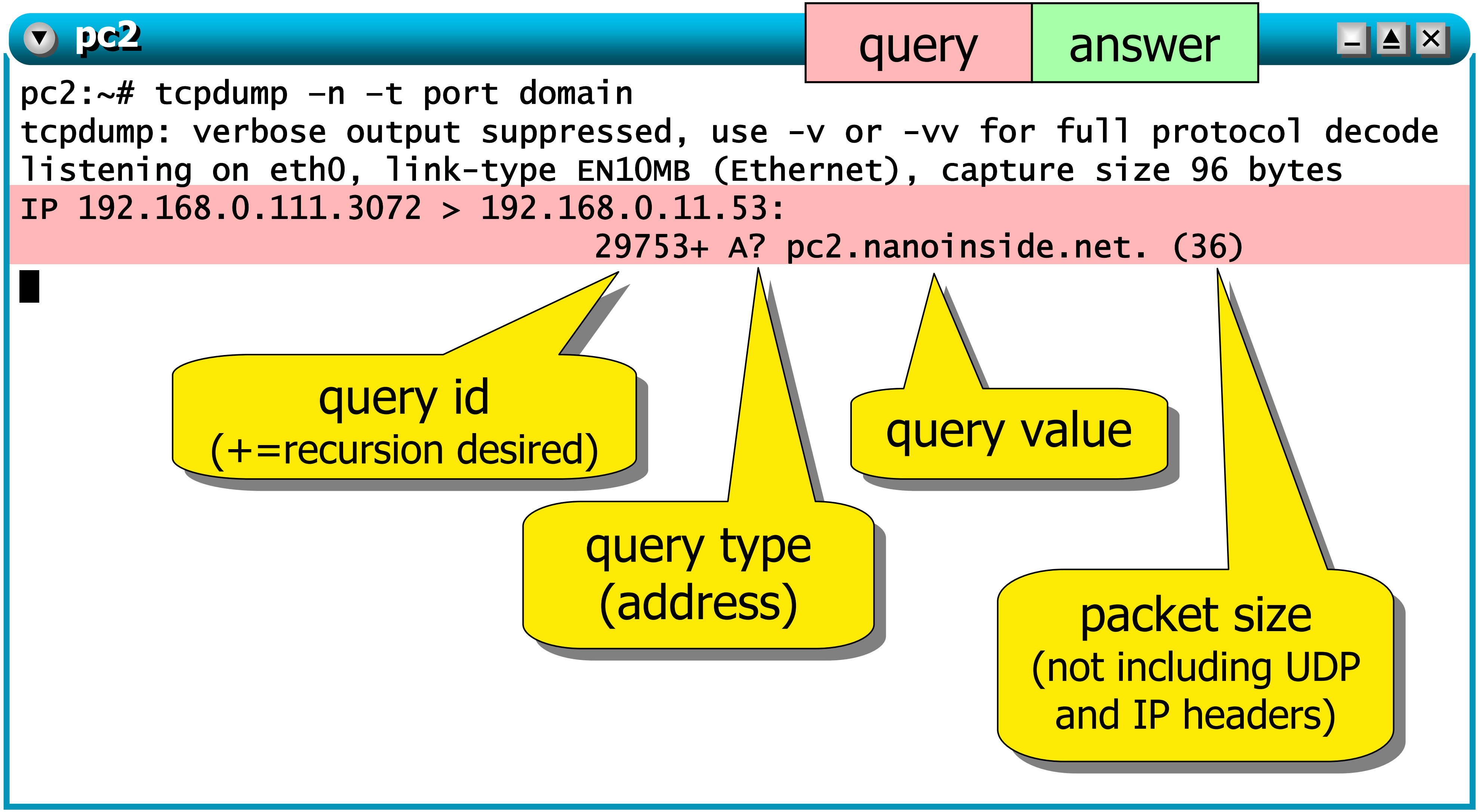
```
pc2:~# tcpdump -n -t port domain
```

ip numbers instead of host names; port numbers instead of service names

no timestamps needed

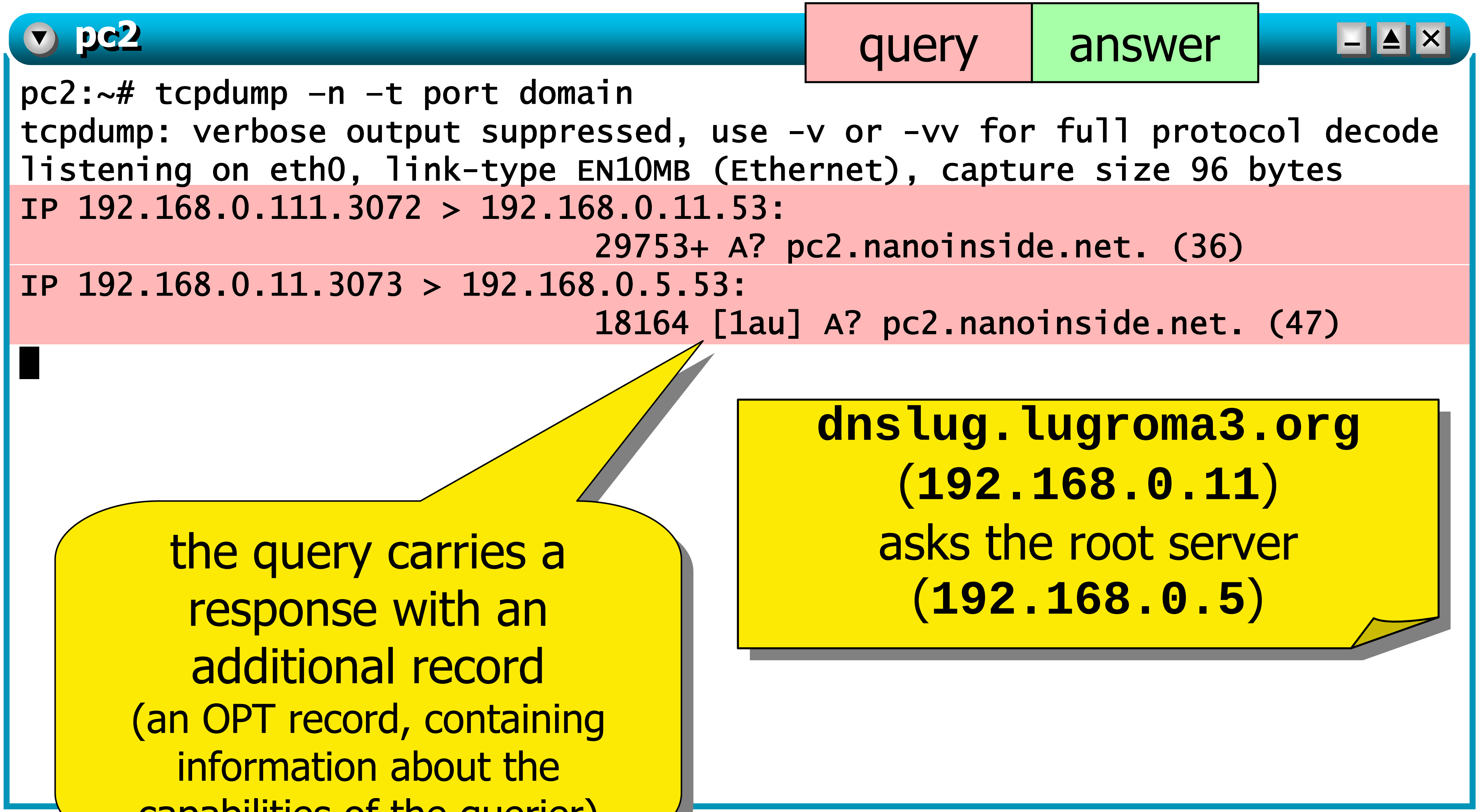
capture packets to/from port “domain” (port 53)

step 3 – the sniffer output



```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
```

step 3 – the sniffer output

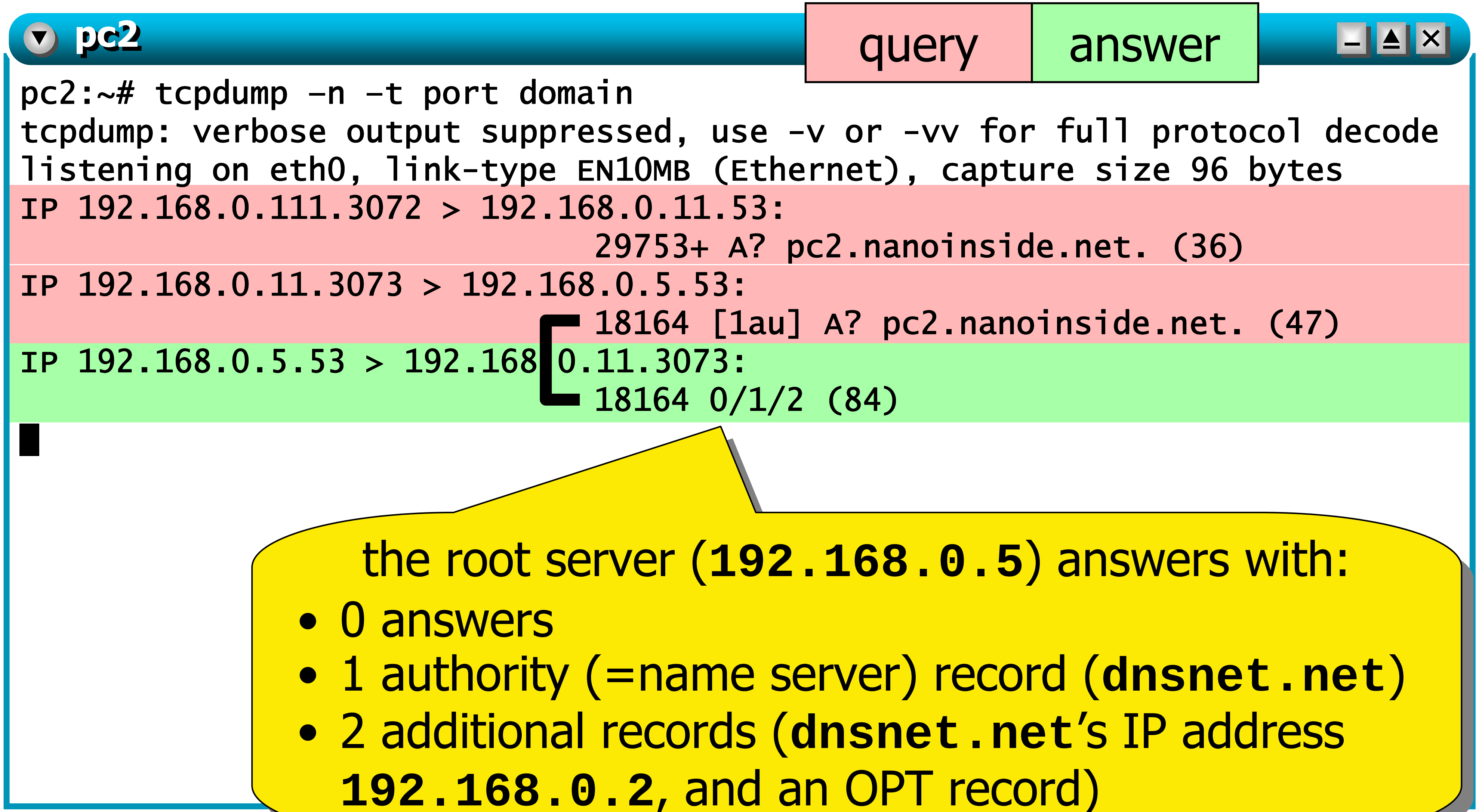


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
IP 192.168.0.11.3073 > 192.168.0.5.53:
                              18164 [1au] A? pc2.nanoinside.net. (47)
```

the query carries a response with an additional record (an OPT record, containing information about the capabilities of the querier)

dnslug.lugroma3.org (192.168.0.11) asks the root server **(192.168.0.5)**

step 3 – the sniffer output

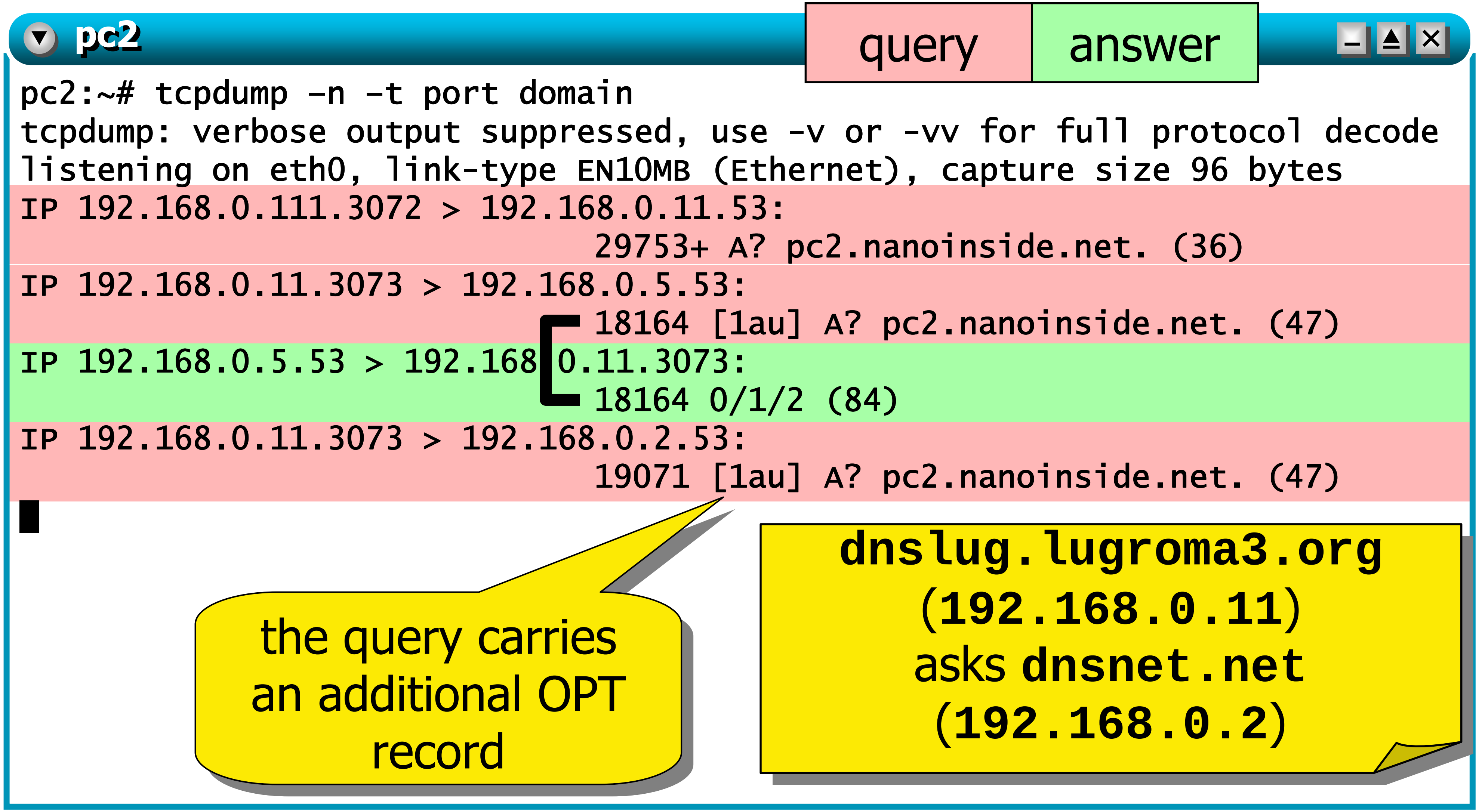


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
IP 192.168.0.11.3073 > 192.168.0.5.53:
                              18164 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.5.53 > 192.168.0.11.3073:
                              18164 0/1/2 (84)
█
```

the root server (**192.168.0.5**) answers with:

- 0 answers
- 1 authority (=name server) record (**dnsnet.net**)
- 2 additional records (**dnsnet.net**'s IP address **192.168.0.2**, and an OPT record)

step 3 – the sniffer output

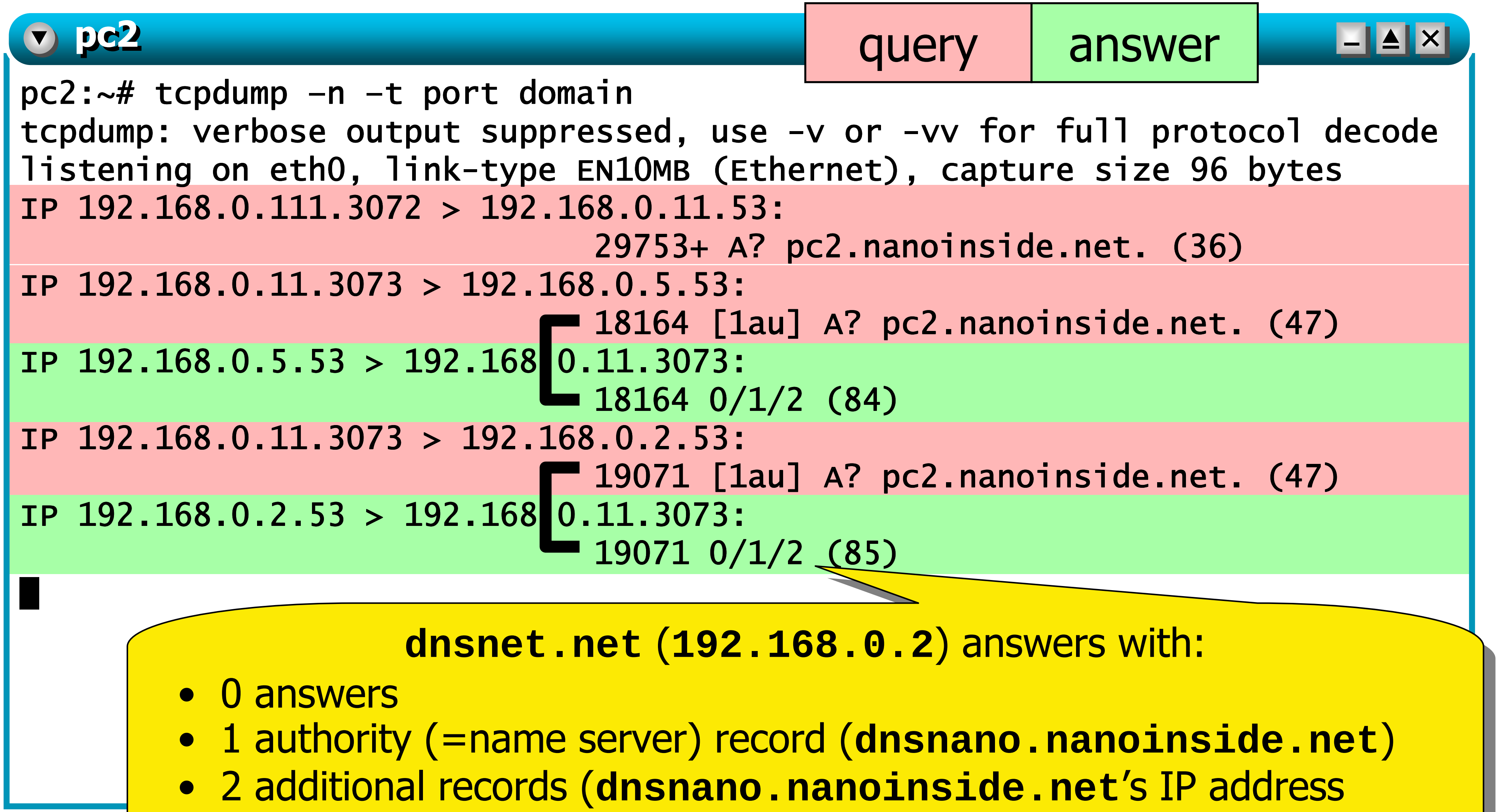


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
IP 192.168.0.11.3073 > 192.168.0.5.53:
                              18164 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.5.53 > 192.168.0.11.3073:
                              18164 0/1/2 (84)
IP 192.168.0.11.3073 > 192.168.0.2.53:
                              19071 [1au] A? pc2.nanoinside.net. (47)
█
```

the query carries an additional OPT record

dnslug.lugroma3.org
(192.168.0.11)
asks **dnsnet.net**
(192.168.0.2)

step 3 – the sniffer output

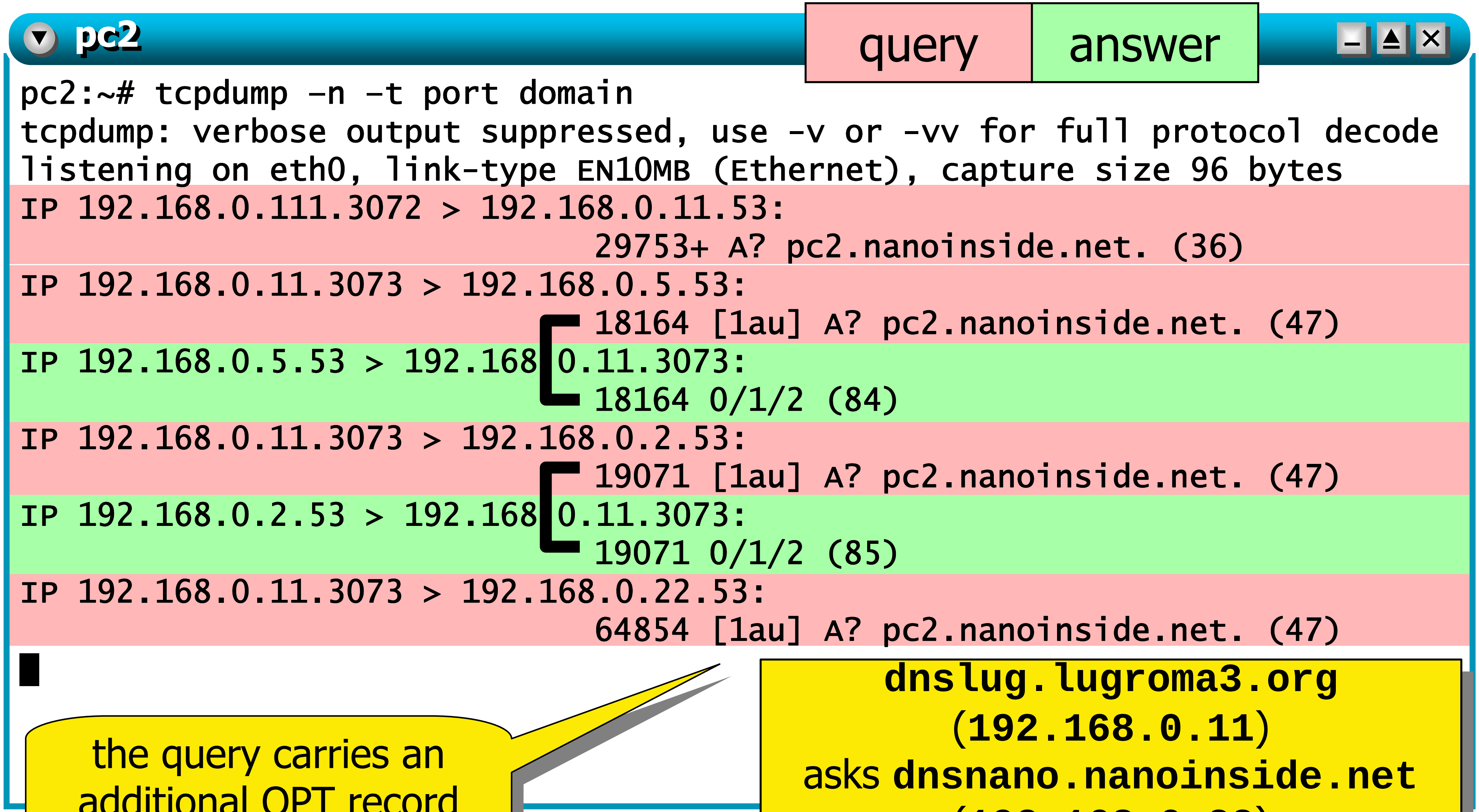


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
IP 192.168.0.11.3073 > 192.168.0.5.53:
                              18164 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.5.53 > 192.168.0.11.3073:
                              18164 0/1/2 (84)
IP 192.168.0.11.3073 > 192.168.0.2.53:
                              19071 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.2.53 > 192.168.0.11.3073:
                              19071 0/1/2 (85)
```

dnsnet.net (**192.168.0.2**) answers with:

- 0 answers
- 1 authority (=name server) record (**dnsnano.nanoinside.net**)
- 2 additional records (**dnsnano.nanoinside.net**'s IP address **192.168.0.22**, and an OPT record)

step 3 – the sniffer output

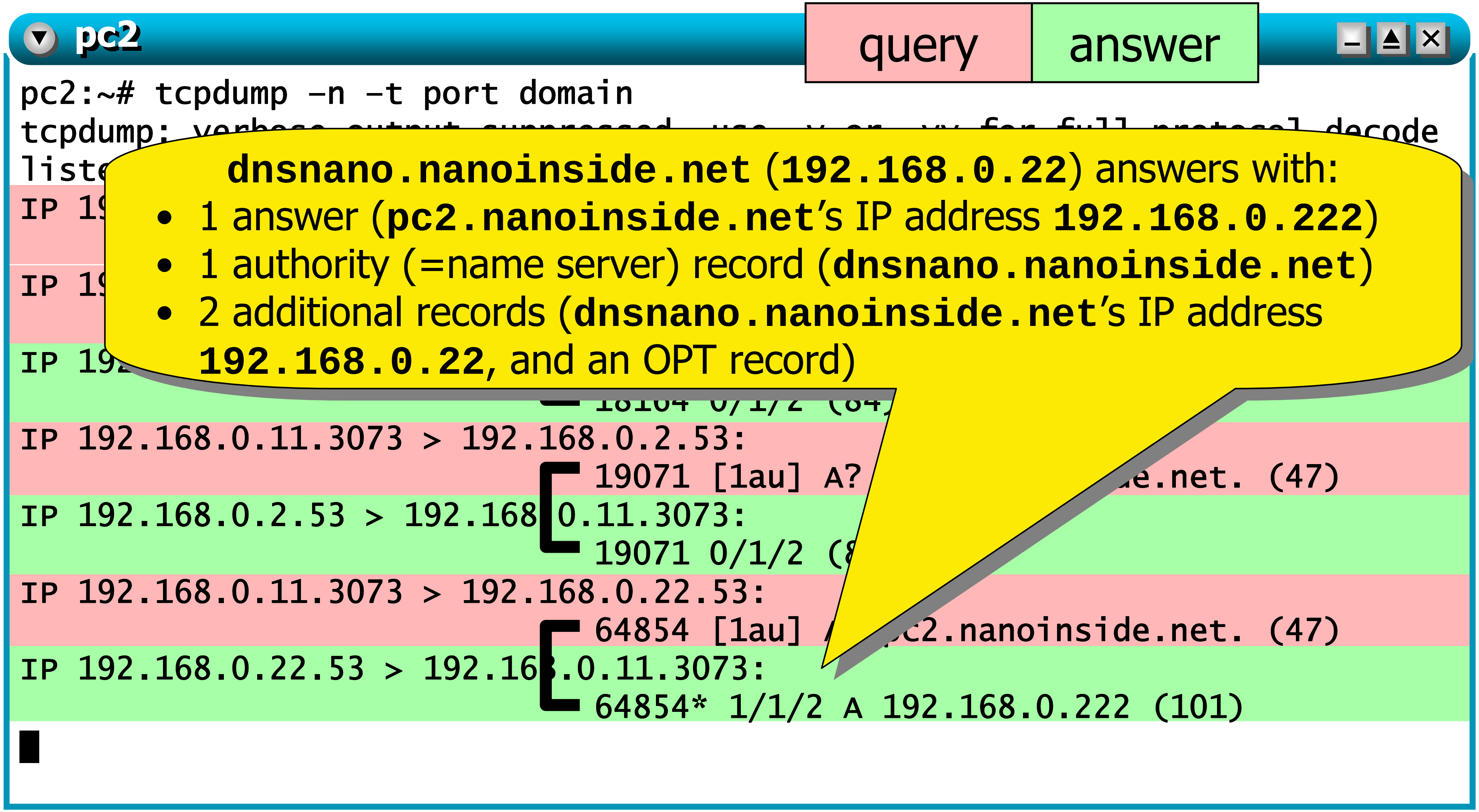


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              29753+ A? pc2.nanoinside.net. (36)
IP 192.168.0.11.3073 > 192.168.0.5.53:
                              18164 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.5.53 > 192.168.0.11.3073:
                              18164 0/1/2 (84)
IP 192.168.0.11.3073 > 192.168.0.2.53:
                              19071 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.2.53 > 192.168.0.11.3073:
                              19071 0/1/2 (85)
IP 192.168.0.11.3073 > 192.168.0.22.53:
                              64854 [1au] A? pc2.nanoinside.net. (47)
█
```

the query carries an additional OPT record

dnslug.lugroma3.org
(192.168.0.11)
asks **dnsnano.nanoinside.net**
(192.168.0.22)

step 3 – the sniffer output

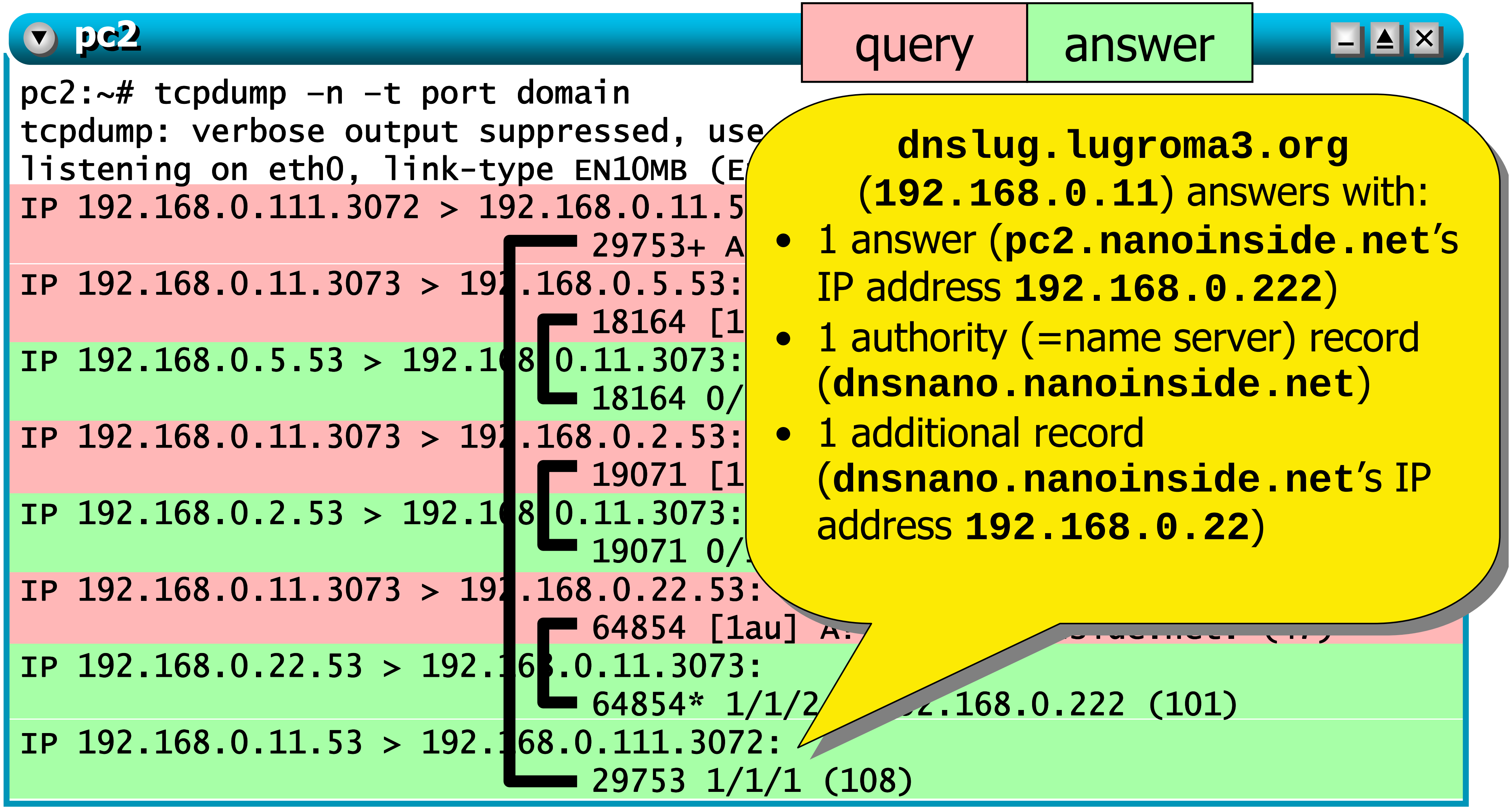


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
liste
IP 19

IP 19

IP 19
                          18164 0/1/2 (84)
IP 192.168.0.11.3073 > 192.168.0.2.53:
                          19071 [1au] A?               e.net. (47)
IP 192.168.0.2.53 > 192.168.0.11.3073:
                          19071 0/1/2 (8
IP 192.168.0.11.3073 > 192.168.0.22.53:
                          64854 [1au] A? pc2.nanoinside.net. (47)
IP 192.168.0.22.53 > 192.168.0.11.3073:
                          64854* 1/1/2 A 192.168.0.222 (101)
█
```

dnsnano.nanoinside.net (**192.168.0.22**) answers with:

- 1 answer (**pc2.nanoinside.net**'s IP address **192.168.0.222**)
- 1 authority (=name server) record (**dnsnano.nanoinside.net**)
- 2 additional records (**dnsnano.nanoinside.net**'s IP address **192.168.0.22**, and an OPT record)

step 3 – the sniffer output

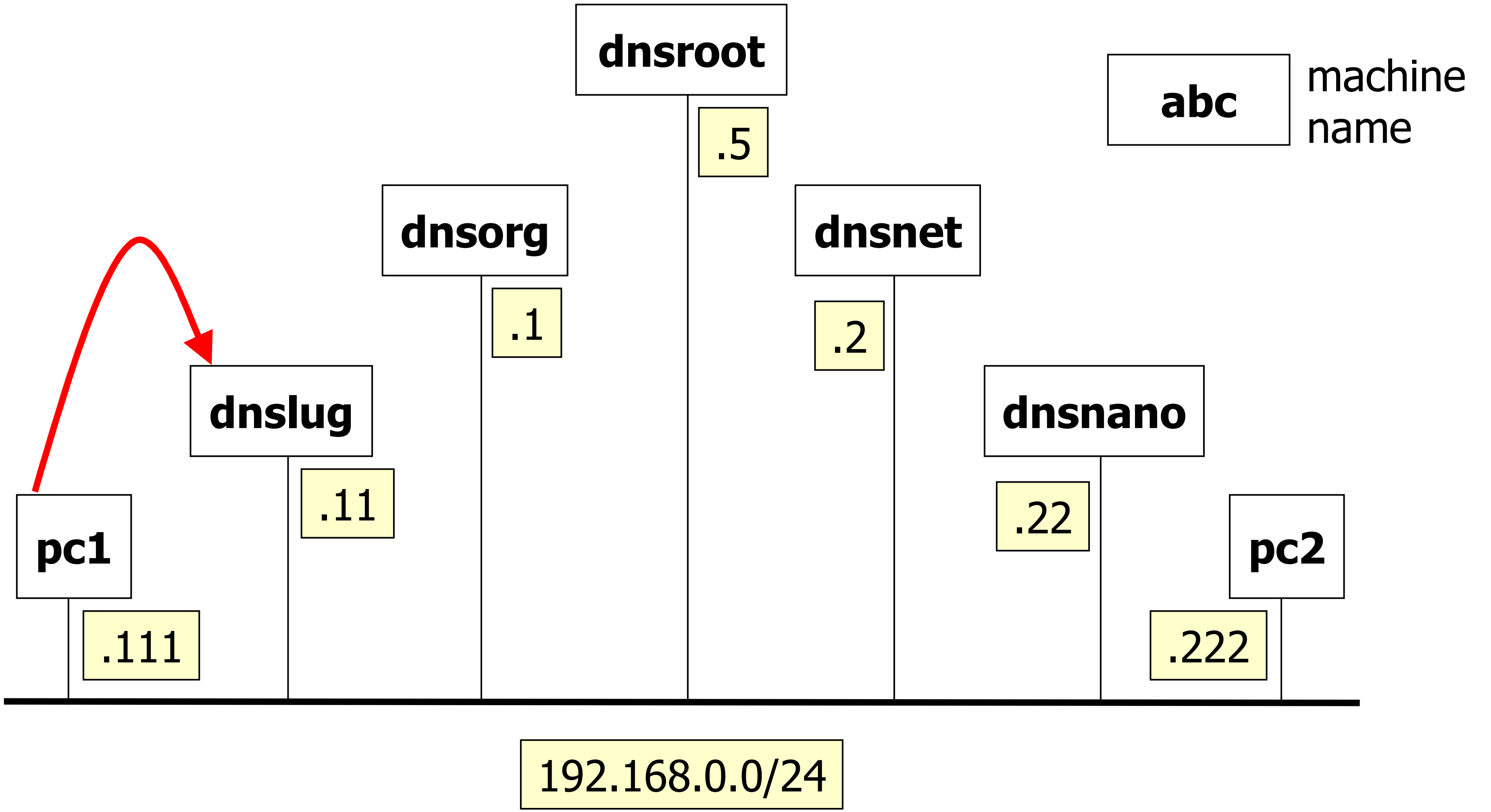


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use
listening on eth0, link-type EN10MB (E
IP 192.168.0.111.3072 > 192.168.0.11.5
                        29753+ A
IP 192.168.0.11.3073 > 192.168.0.5.53:
                        18164 [1
IP 192.168.0.5.53 > 192.168.0.11.3073:
                        18164 0/
IP 192.168.0.11.3073 > 192.168.0.2.53:
                        19071 [1
IP 192.168.0.2.53 > 192.168.0.11.3073:
                        19071 0/
IP 192.168.0.11.3073 > 192.168.0.22.53:
                        64854 [1au] A
IP 192.168.0.22.53 > 192.168.0.11.3073:
                        64854* 1/1/2     2.168.0.222 (101)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                        29753 1/1/1 (108)
```

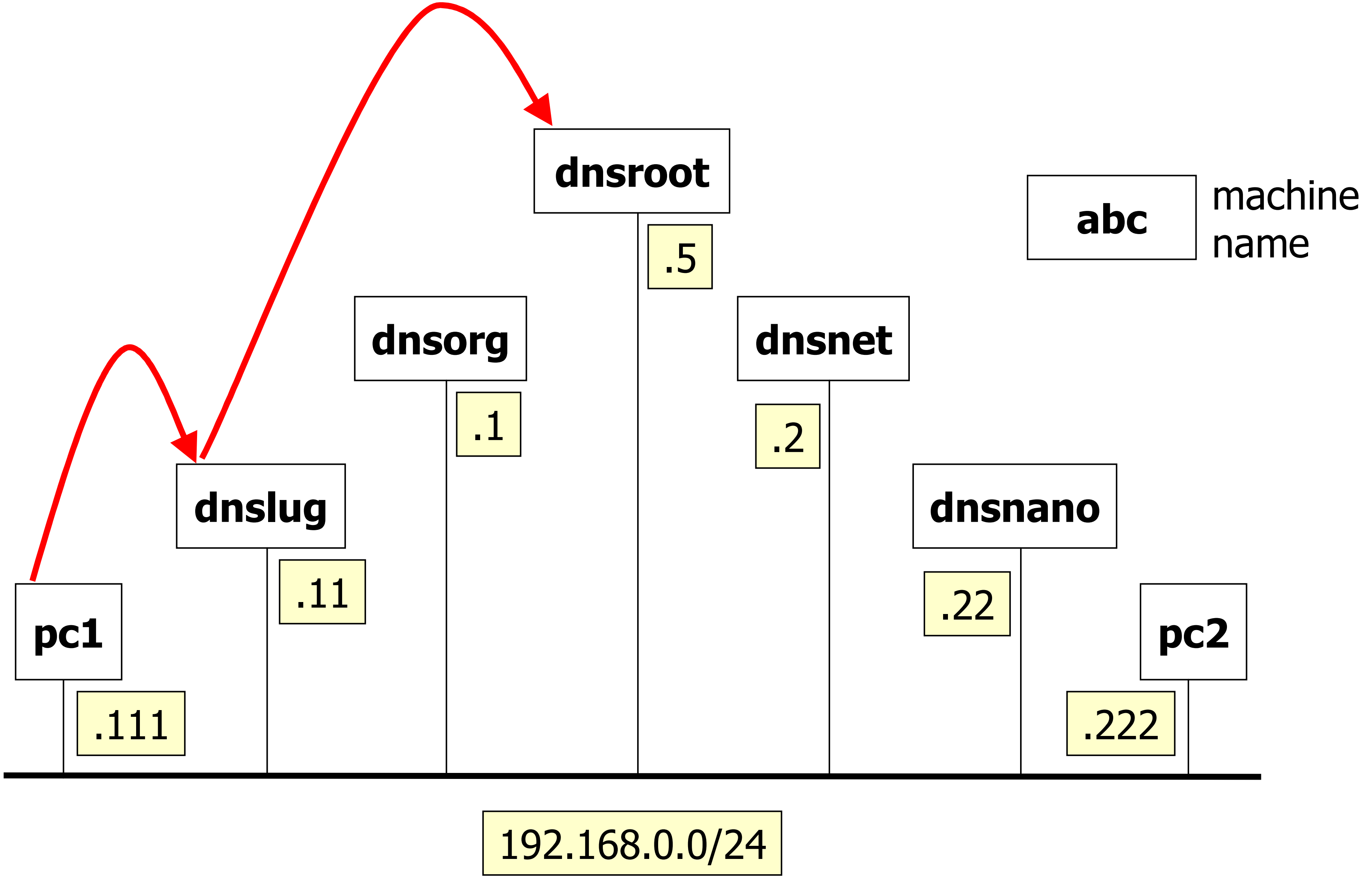
dnslug.lugroma3.org (**192.168.0.11**) answers with:

- 1 answer (**pc2.nanoinside.net**'s IP address **192.168.0.222**)
- 1 authority (=name server) record (**dnsnano.nanoinside.net**)
- 1 additional record (**dnsnano.nanoinside.net**'s IP address **192.168.0.22**)

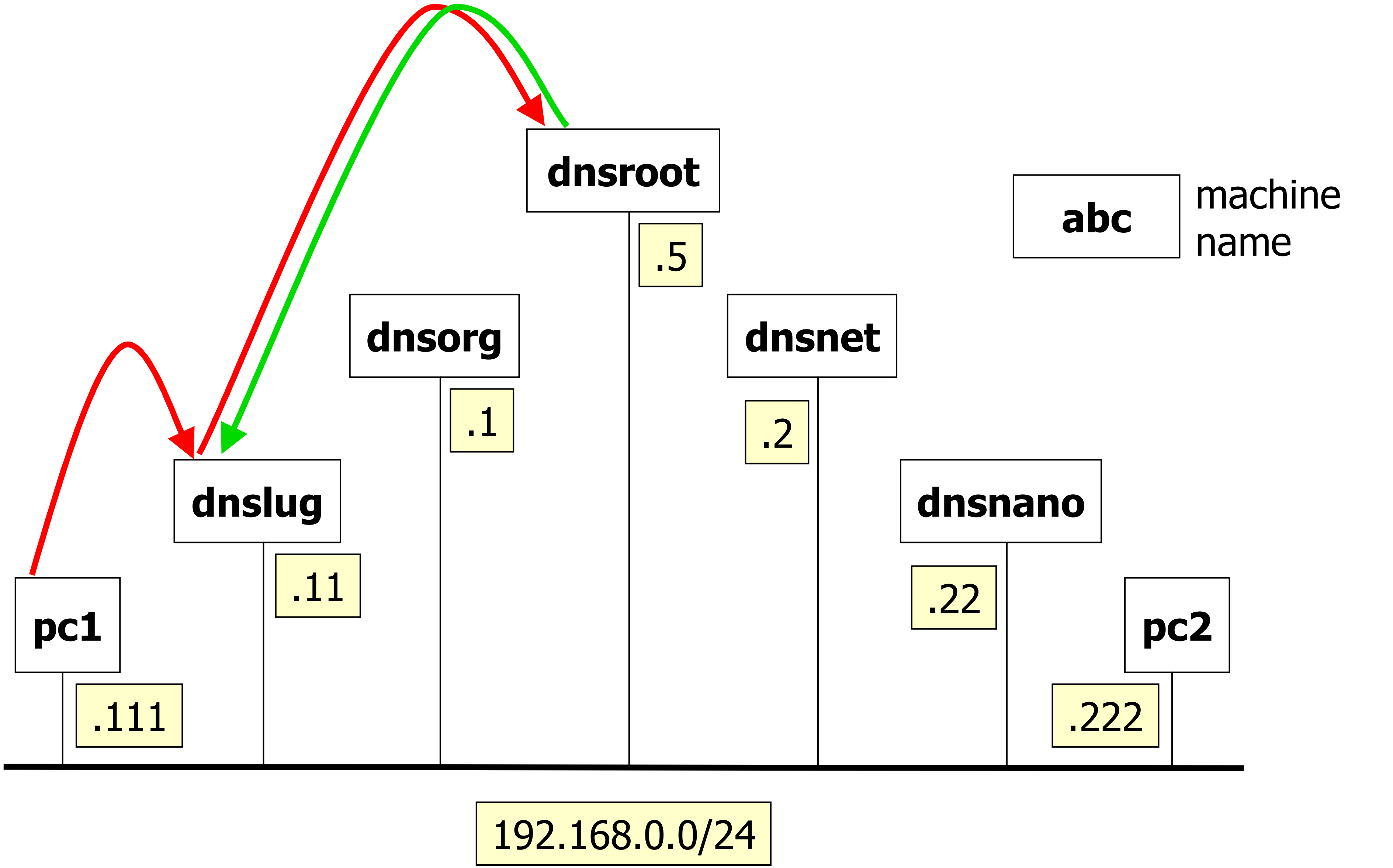
step 3 – exchanged messages



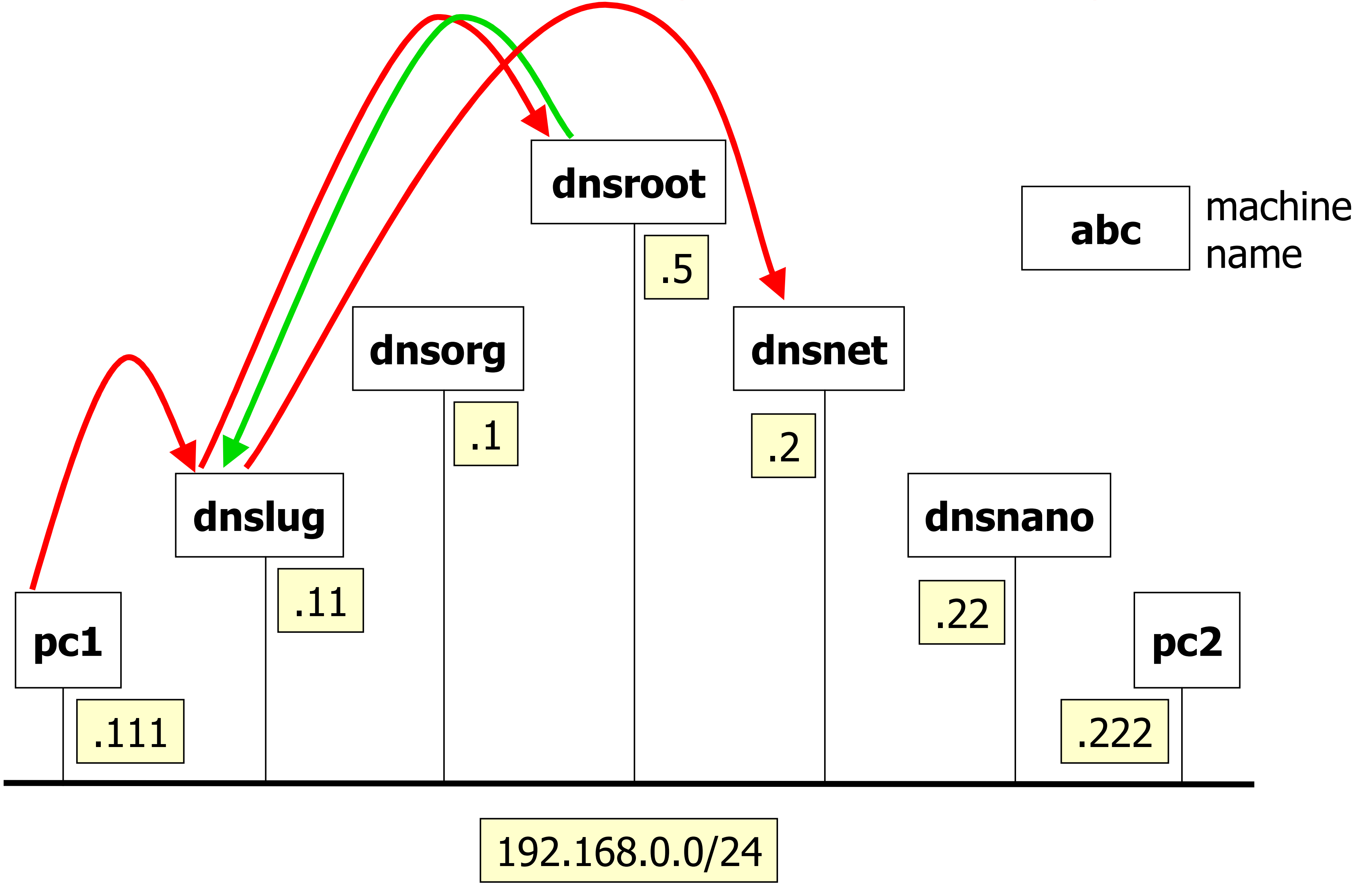
step 3 – exchanged messages



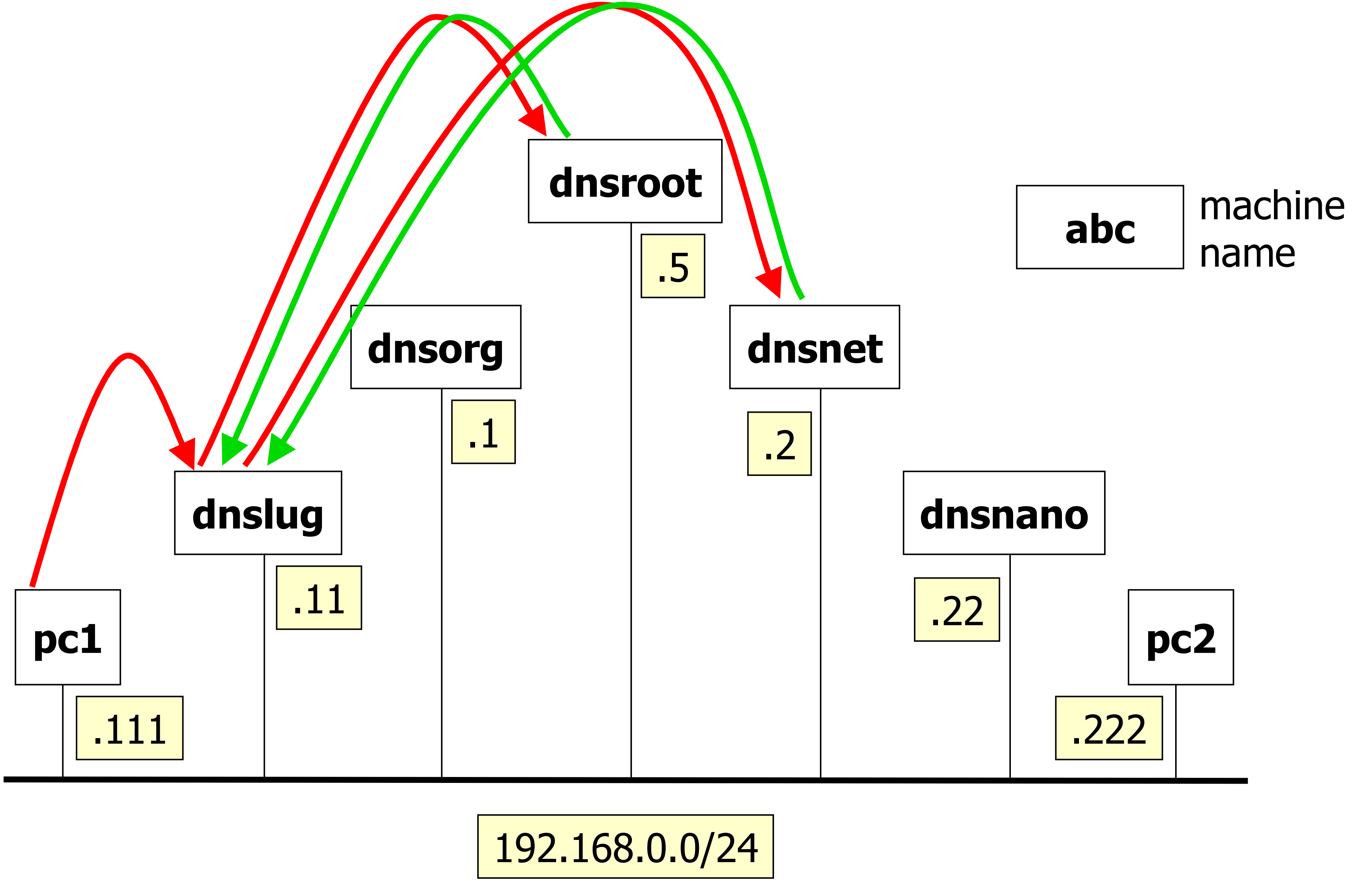
step 3 – exchanged messages



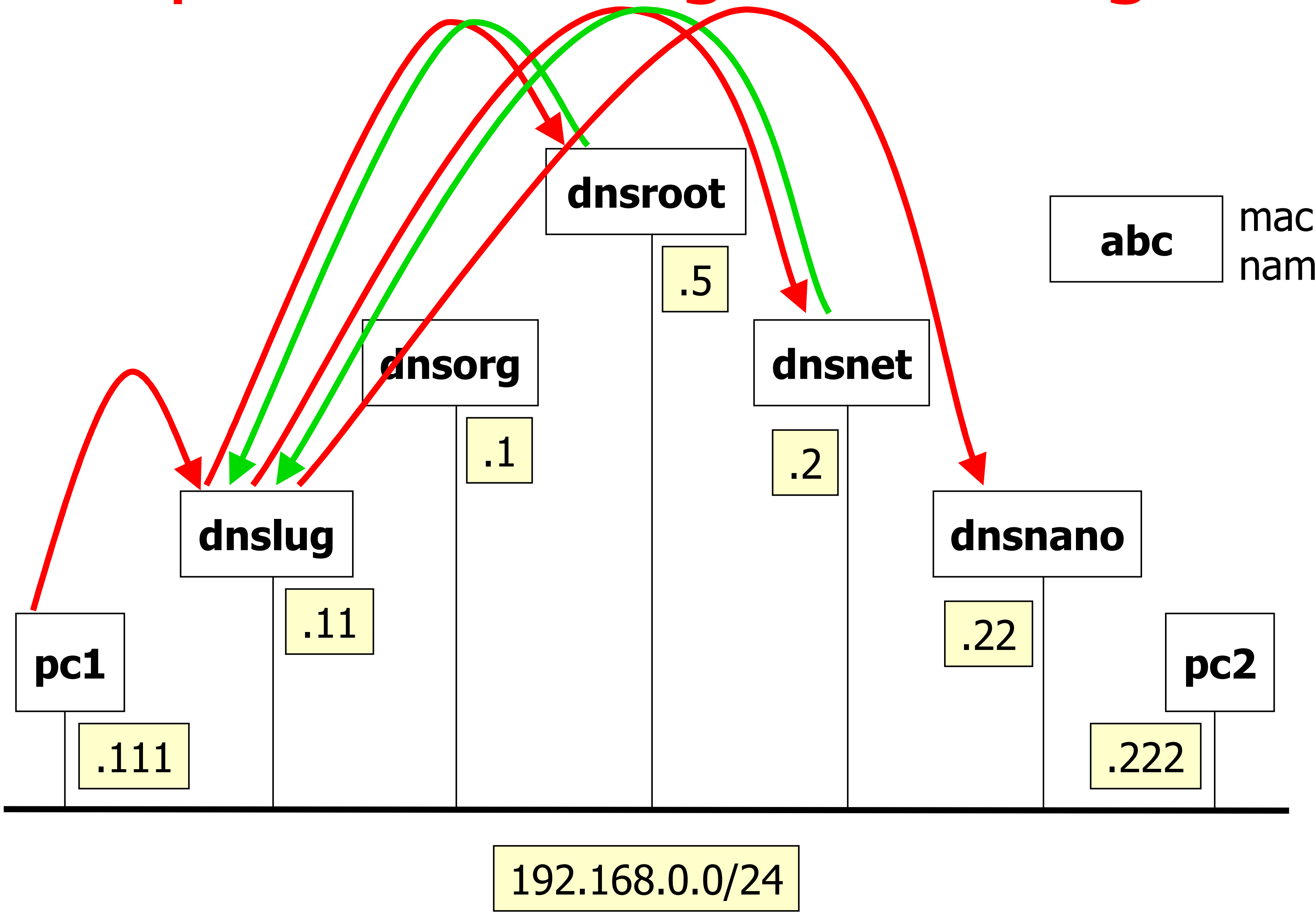
step 3 – exchanged messages



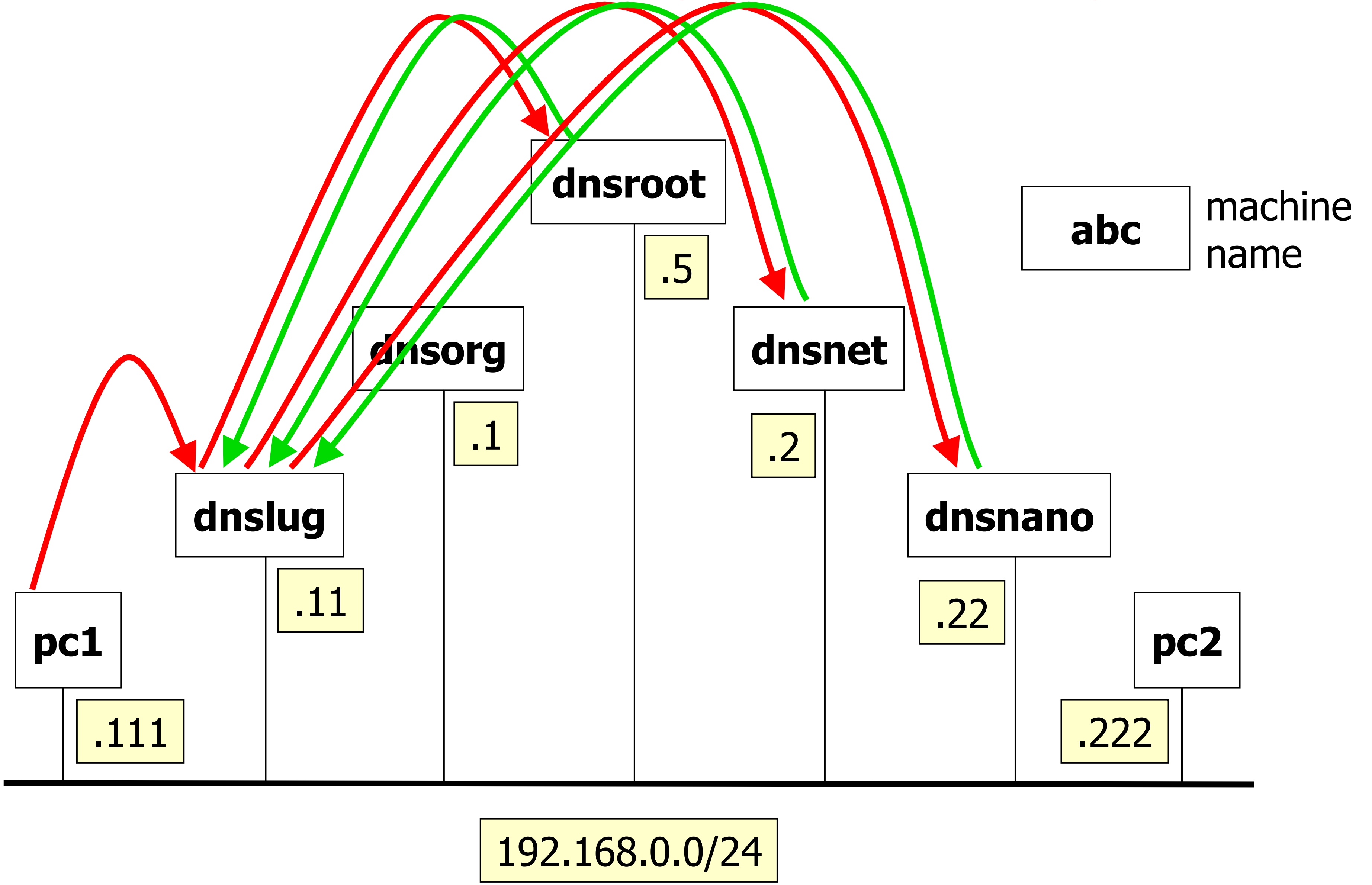
step 3 – exchanged messages



step 3 – exchanged messages

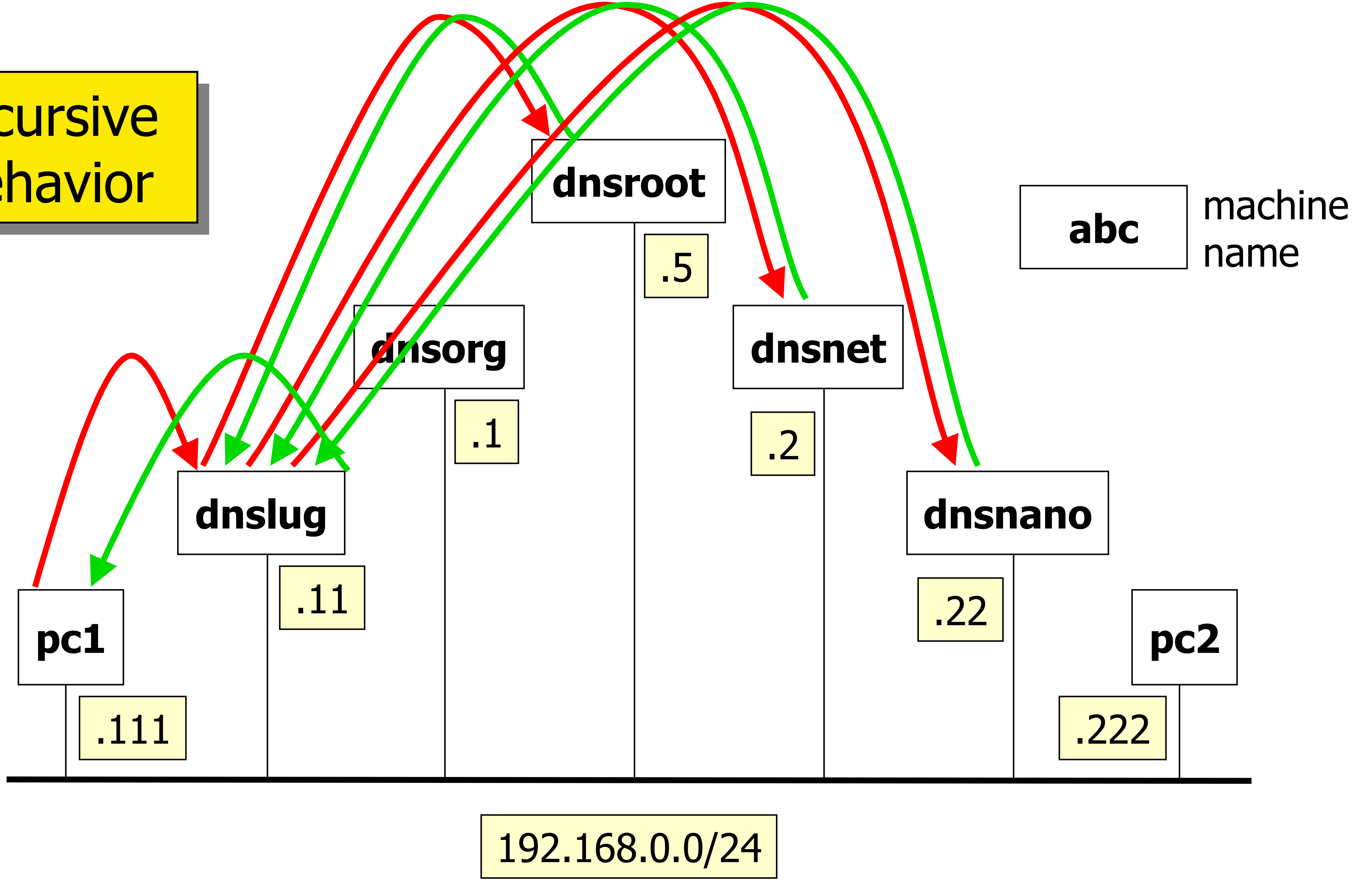


step 3 – exchanged messages

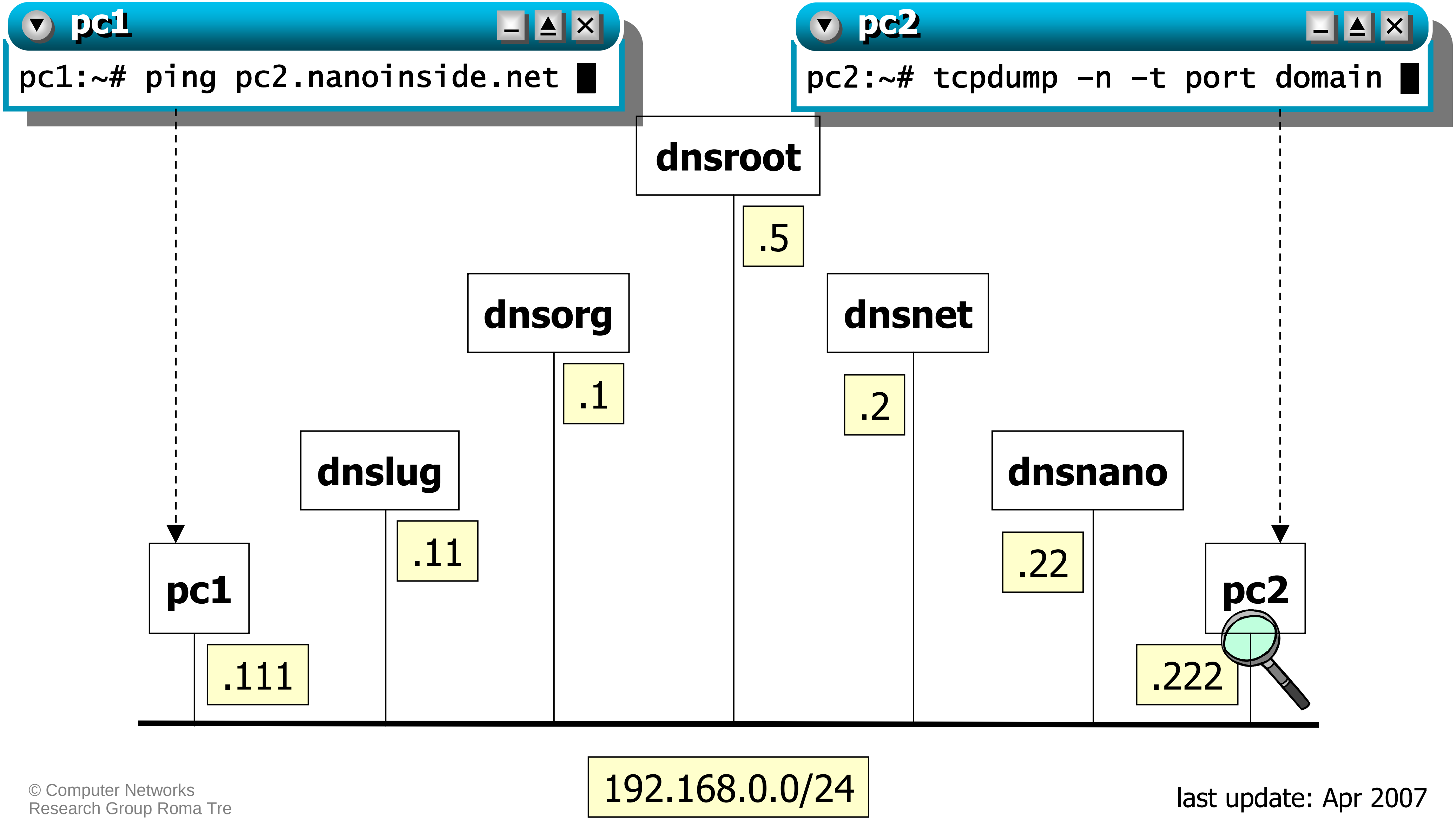


step 3 – exchanged messages

recursive behavior



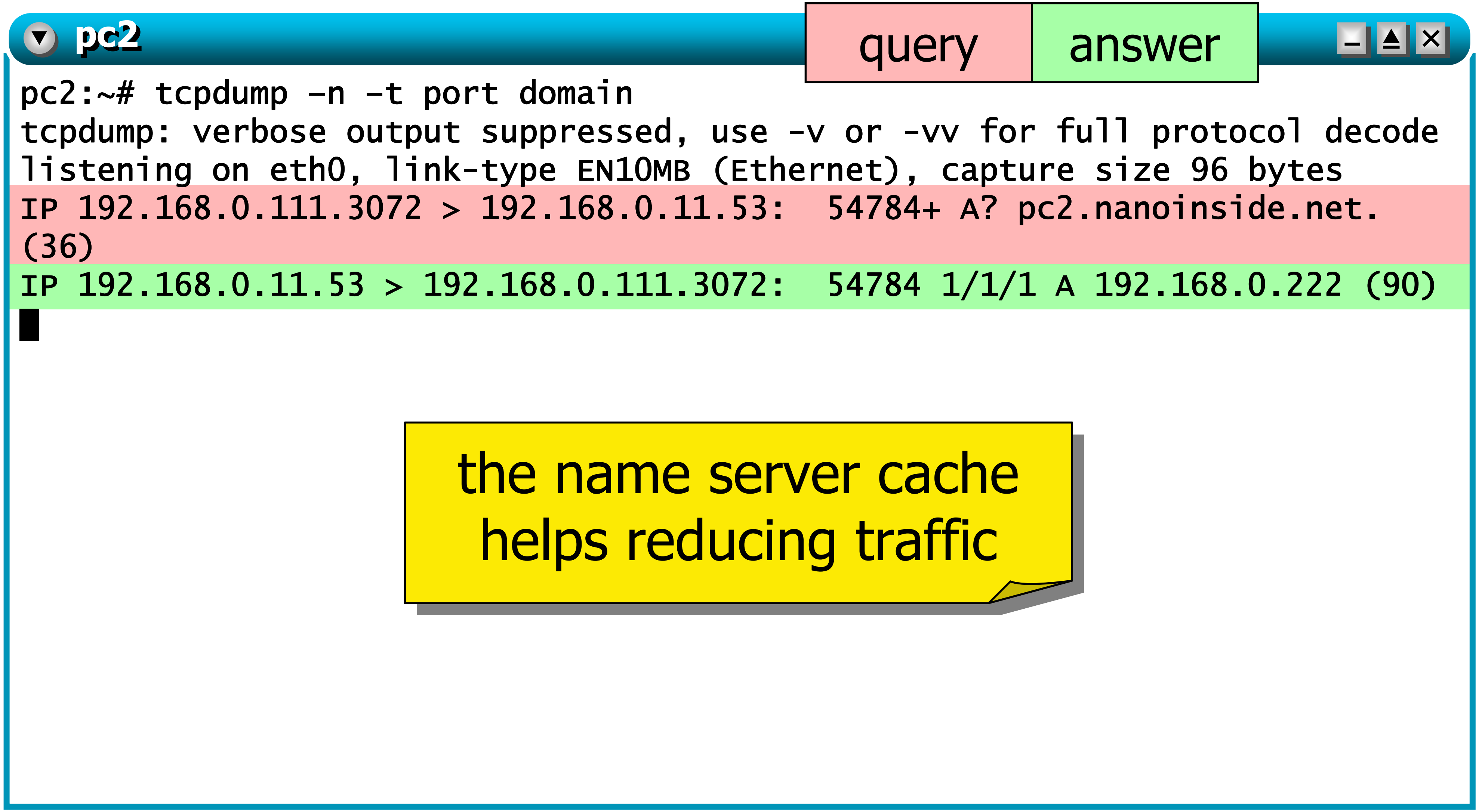
step 4 – repeating the experiment



```
pc1:~# ping pc2.nanoinside.net
```

```
pc2:~# tcpdump -n -t port domain
```

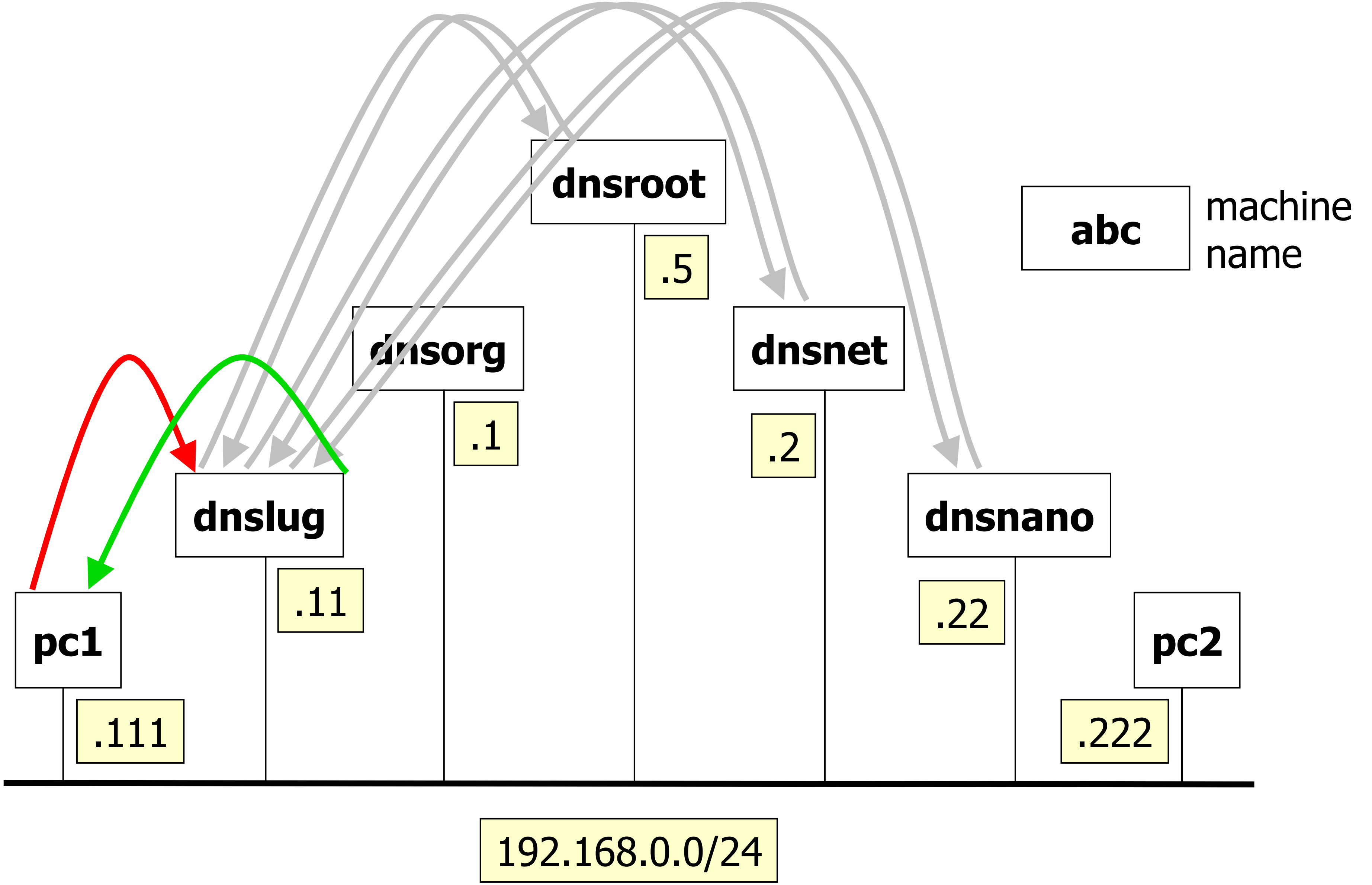
step 4 – repeating the experiment



```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:  54784+ A? pc2.nanoinside.net.
(36)
IP 192.168.0.11.53 > 192.168.0.111.3072:  54784 1/1/1 A 192.168.0.222 (90)
█
```

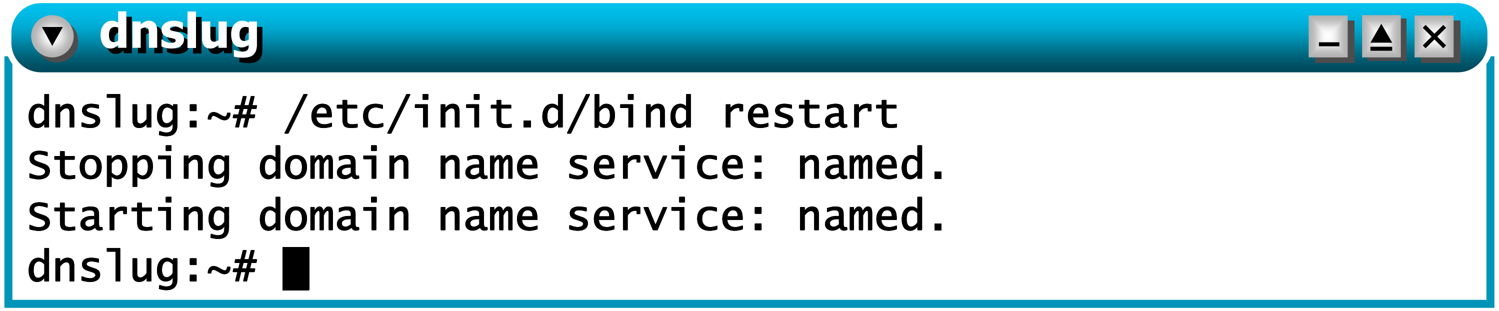
the name server cache helps reducing traffic

step 4 – repeating the experiment



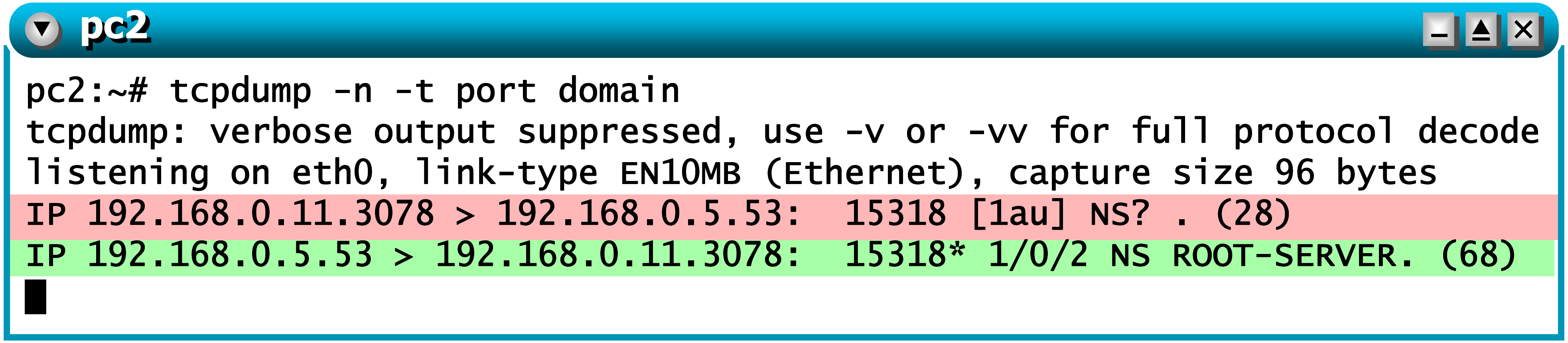
step 5 – restarting the name server

- the restart operation cleans up caches
  - a new client query triggers the complete sequence of iterative queries



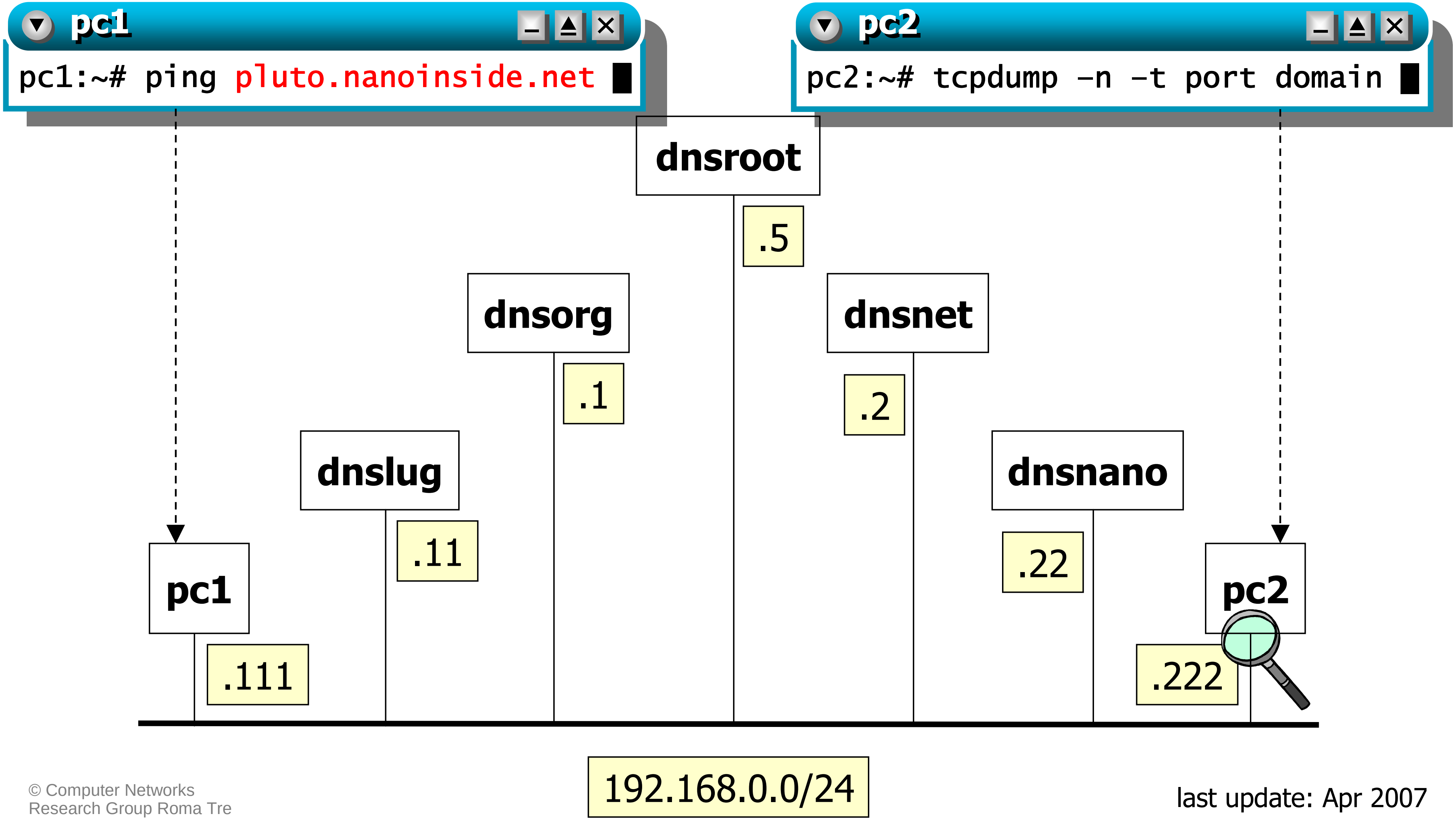
```
dnslug:~# /etc/init.d/bind restart
Stopping domain name service: named.
Starting domain name service: named.
dnslug:~# █
```

- upon startup, the name server checks its root server configuration



```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.11.3078 > 192.168.0.5.53:  15318 [1au] NS? . (28)
IP 192.168.0.5.53 > 192.168.0.11.3078:  15318* 1/0/2 NS ROOT-SERVER. (68)
█
```

step 6 – non-existent target

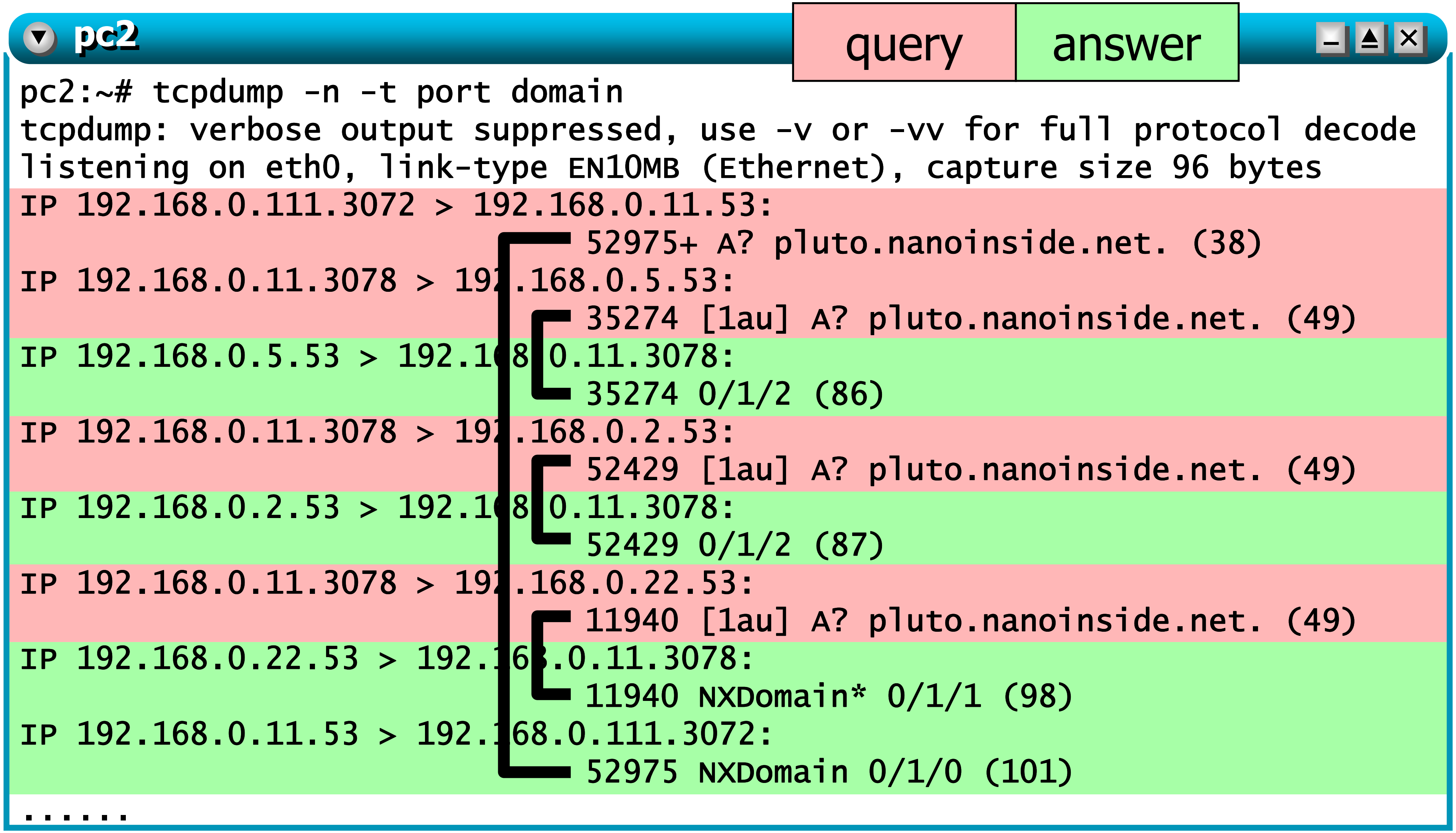


```
pc1:~# ping pluto.nanoinside.net
```

```
pc2:~# tcpdump -n -t port domain
```

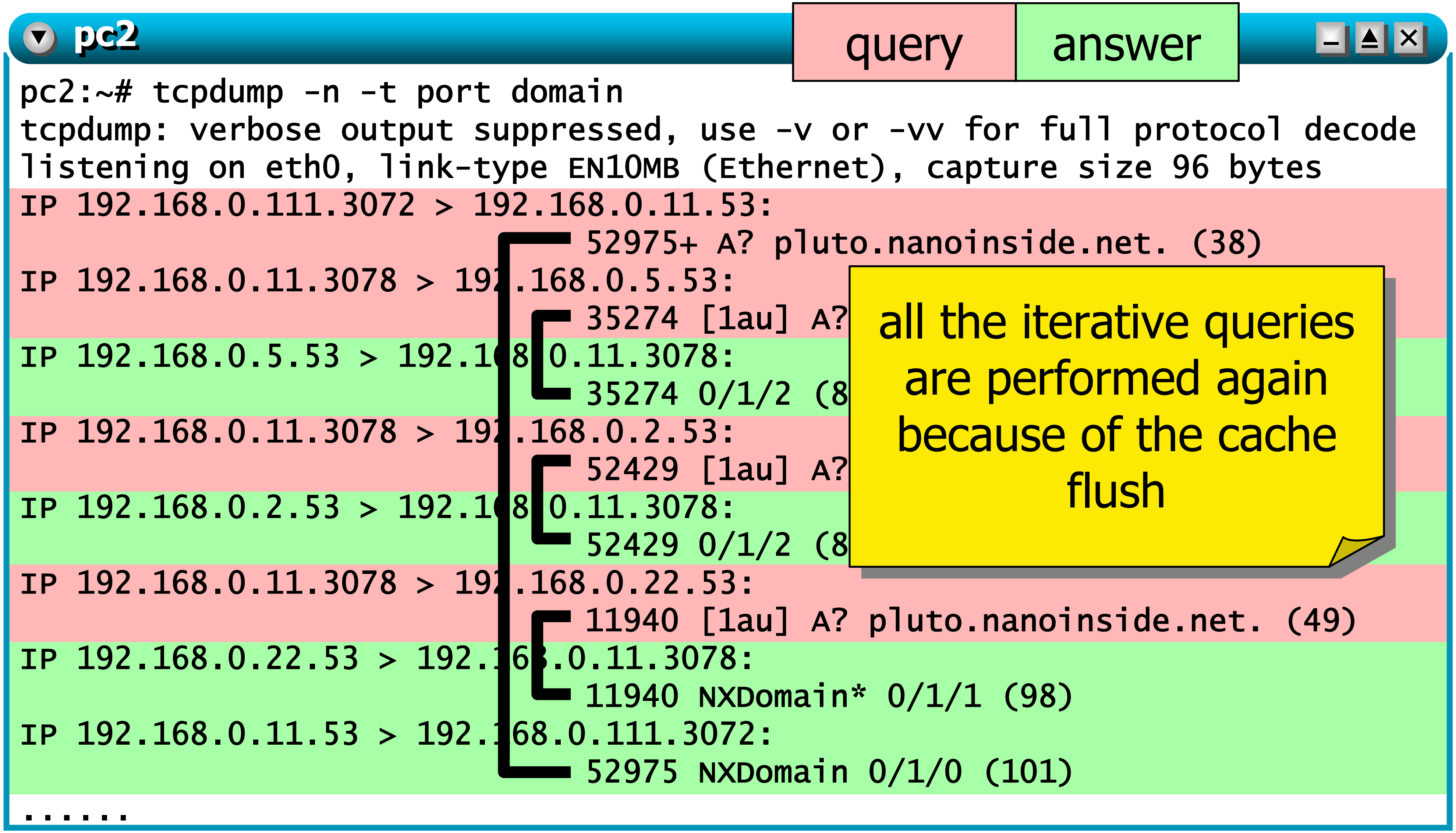
last update: Apr 2007

step 6 – non-existent target



```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              52975+ A? pluto.nanoinside.net. (38)
IP 192.168.0.11.3078 > 192.168.0.5.53:
                              35274 [1au] A? pluto.nanoinside.net. (49)
IP 192.168.0.5.53 > 192.168.0.11.3078:
                              35274 0/1/2 (86)
IP 192.168.0.11.3078 > 192.168.0.2.53:
                              52429 [1au] A? pluto.nanoinside.net. (49)
IP 192.168.0.2.53 > 192.168.0.11.3078:
                              52429 0/1/2 (87)
IP 192.168.0.11.3078 > 192.168.0.22.53:
                              11940 [1au] A? pluto.nanoinside.net. (49)
IP 192.168.0.22.53 > 192.168.0.11.3078:
                              11940 NXDomain* 0/1/1 (98)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                              52975 NXDomain 0/1/0 (101)
......
```

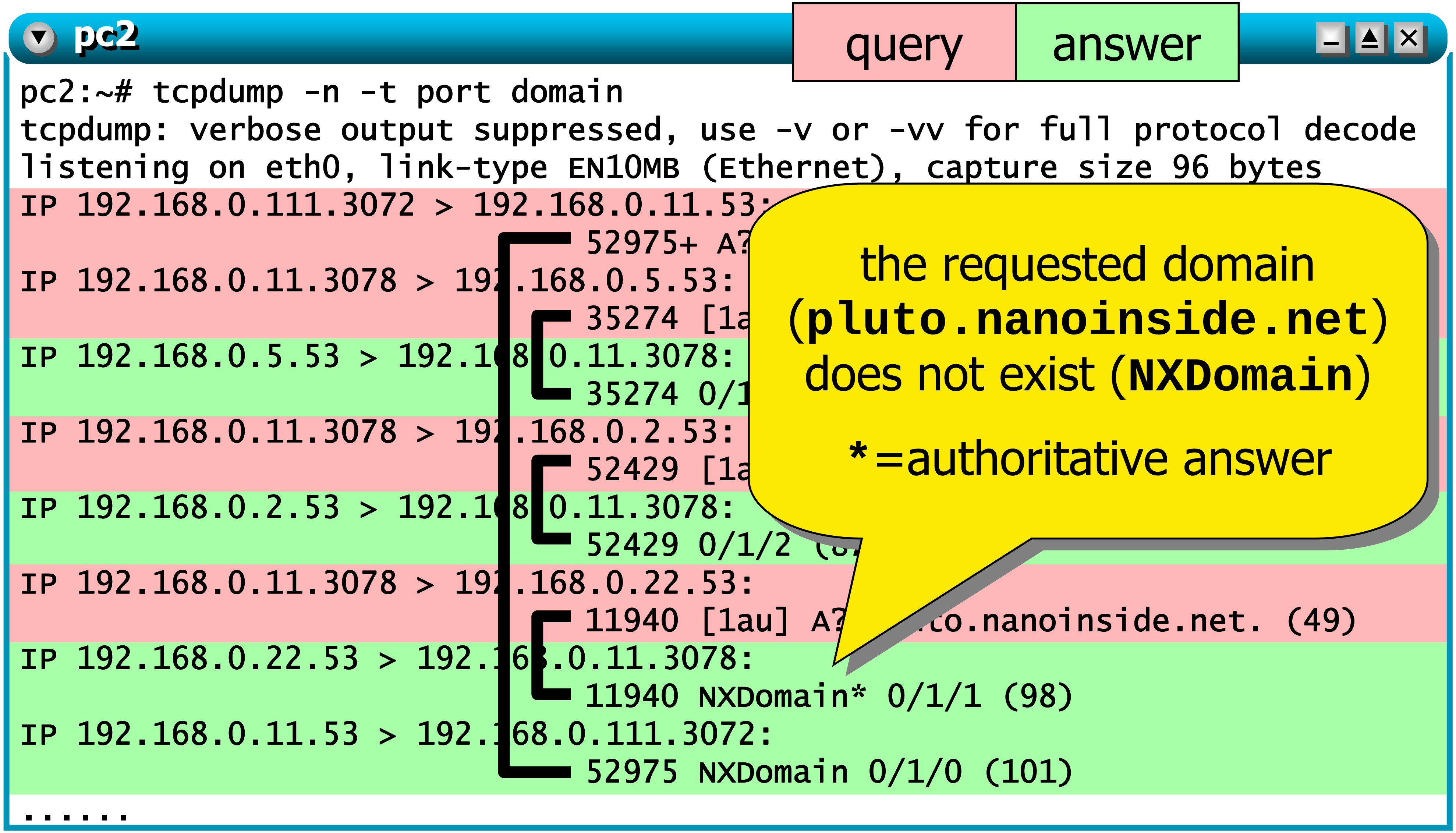
step 6 – non-existent target



```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              52975+ A? pluto.nanoinside.net. (38)
IP 192.168.0.11.3078 > 192.168.0.5.53:
                              35274 [1au] A?
IP 192.168.0.5.53 > 192.168.0.11.3078:
                              35274 0/1/2 (8
IP 192.168.0.11.3078 > 192.168.0.2.53:
                              52429 [1au] A?
IP 192.168.0.2.53 > 192.168.0.11.3078:
                              52429 0/1/2 (8
IP 192.168.0.11.3078 > 192.168.0.22.53:
                              11940 [1au] A? pluto.nanoinside.net. (49)
IP 192.168.0.22.53 > 192.168.0.11.3078:
                              11940 NXDomain* 0/1/1 (98)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                              52975 NXDomain 0/1/0 (101)
......
```

all the iterative queries are performed again because of the cache flush

step 6 – non-existent target

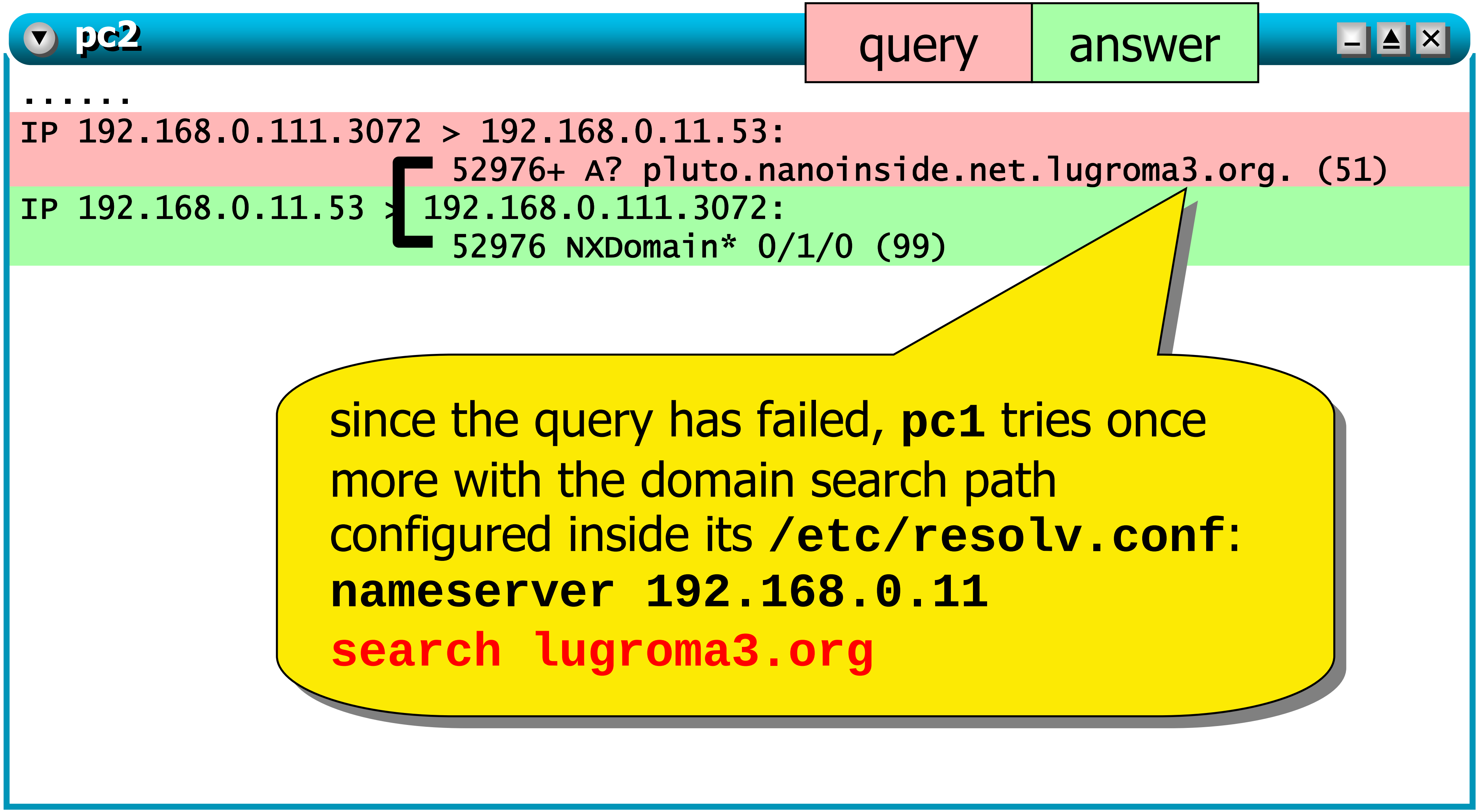


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                              52975+ A?
IP 192.168.0.11.3078 > 192.168.0.5.53:
                              35274 [1a
IP 192.168.0.5.53 > 192.168.0.11.3078:
                              35274 0/1
IP 192.168.0.11.3078 > 192.168.0.2.53:
                              52429 [1a
IP 192.168.0.2.53 > 192.168.0.11.3078:
                              52429 0/1/2 (
IP 192.168.0.11.3078 > 192.168.0.22.53:
                              11940 [1au] A?   to.nanoinside.net. (49)
IP 192.168.0.22.53 > 192.168.0.11.3078:
                              11940 NXDomain* 0/1/1 (98)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                              52975 NXDomain 0/1/0 (101)
......
```

the requested domain (**pluto.nanoinside.net**) does not exist (**NXDomain**)

*****=authoritative answer

step 6 – non-existent target (cont'd)

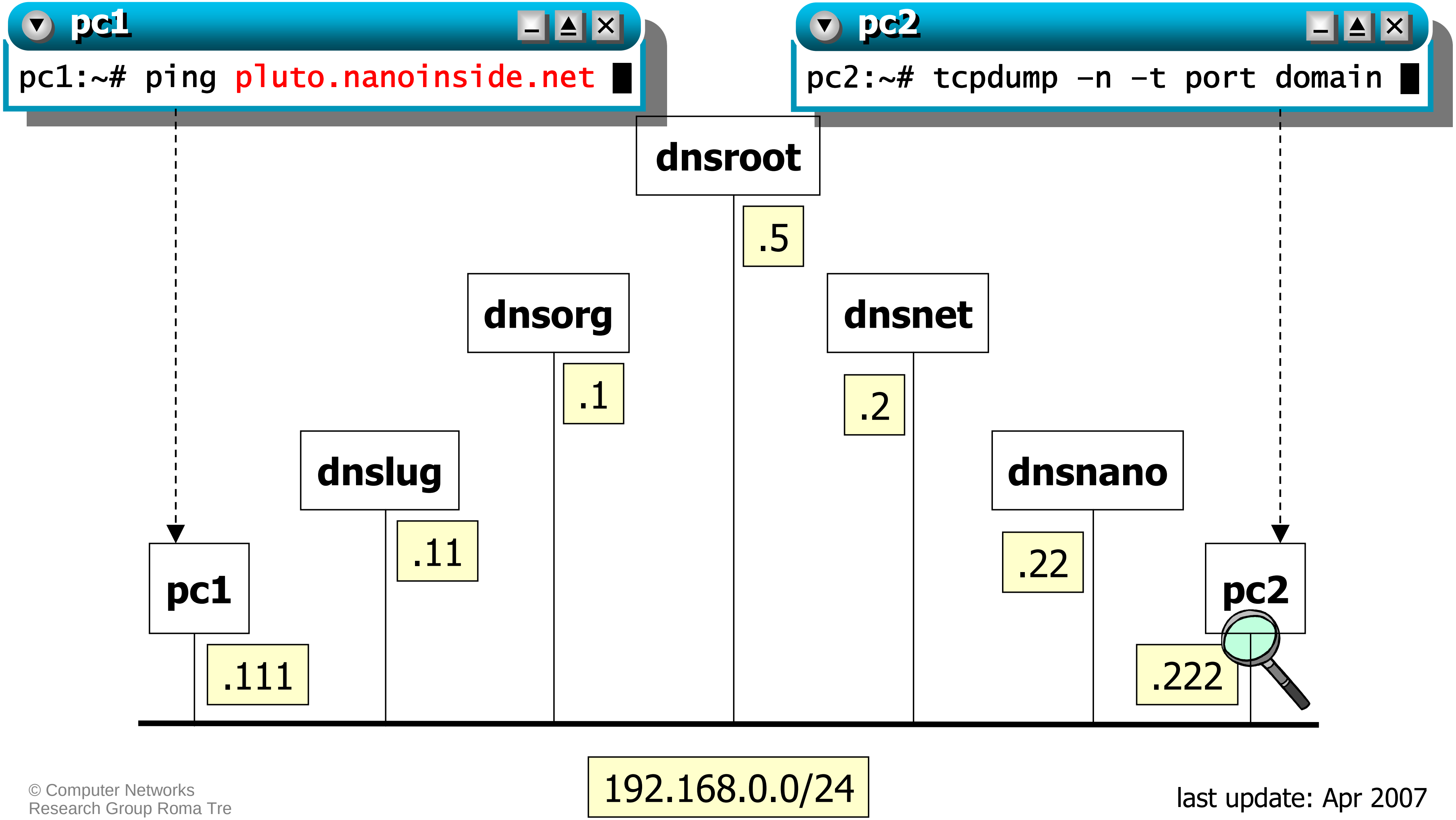


```
......
IP 192.168.0.111.3072 > 192.168.0.11.53:
                          52976+ A? pluto.nanoinside.net.lugroma3.org. (51)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                          52976 NXDomain* 0/1/0 (99)
```

since the query has failed, **`pc1`** tries once more with the domain search path configured inside its **`/etc/resolv.conf`**:

```
nameserver 192.168.0.11
search lugroma3.org
```

step 6 – repeating the experiment

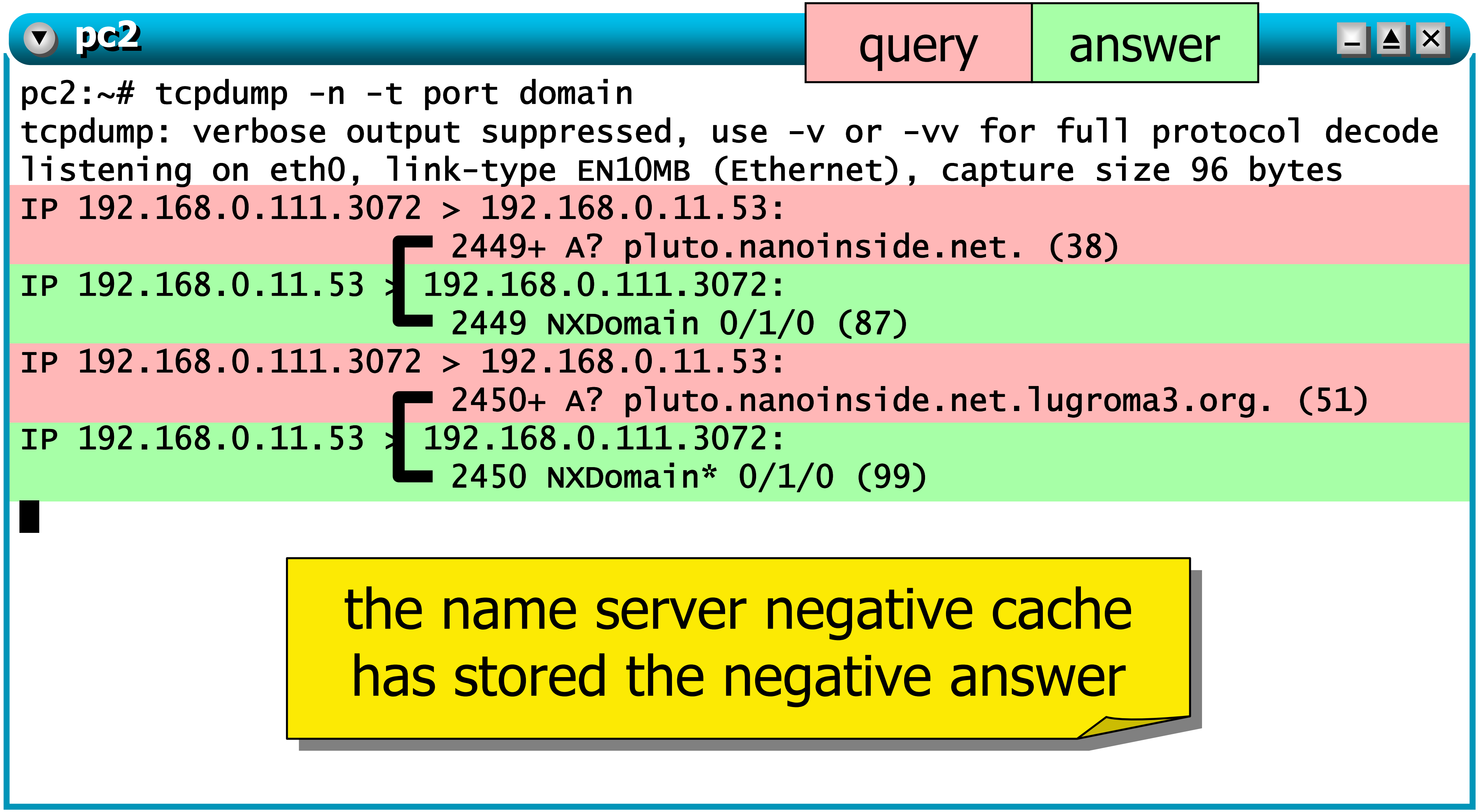


```
pc1:~# ping pluto.nanoinside.net
```

```
pc2:~# tcpdump -n -t port domain
```

last update: Apr 2007

step 6 – repeating the experiment

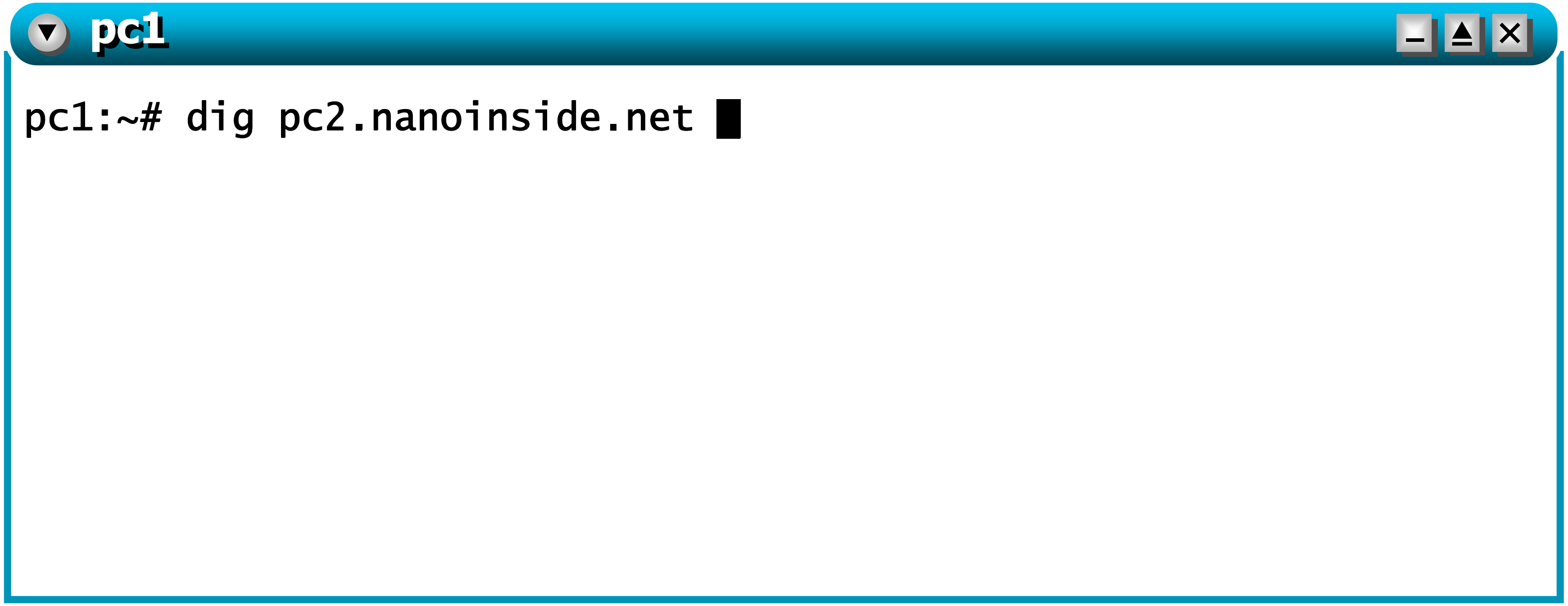


```
pc2:~# tcpdump -n -t port domain
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
IP 192.168.0.111.3072 > 192.168.0.11.53:
                      2449+ A? pluto.nanoinside.net. (38)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                      2449 NXDomain 0/1/0 (87)
IP 192.168.0.111.3072 > 192.168.0.11.53:
                      2450+ A? pluto.nanoinside.net.lugroma3.org. (51)
IP 192.168.0.11.53 > 192.168.0.111.3072:
                      2450 NXDomain* 0/1/0 (99)
```

the name server negative cache has stored the negative answer

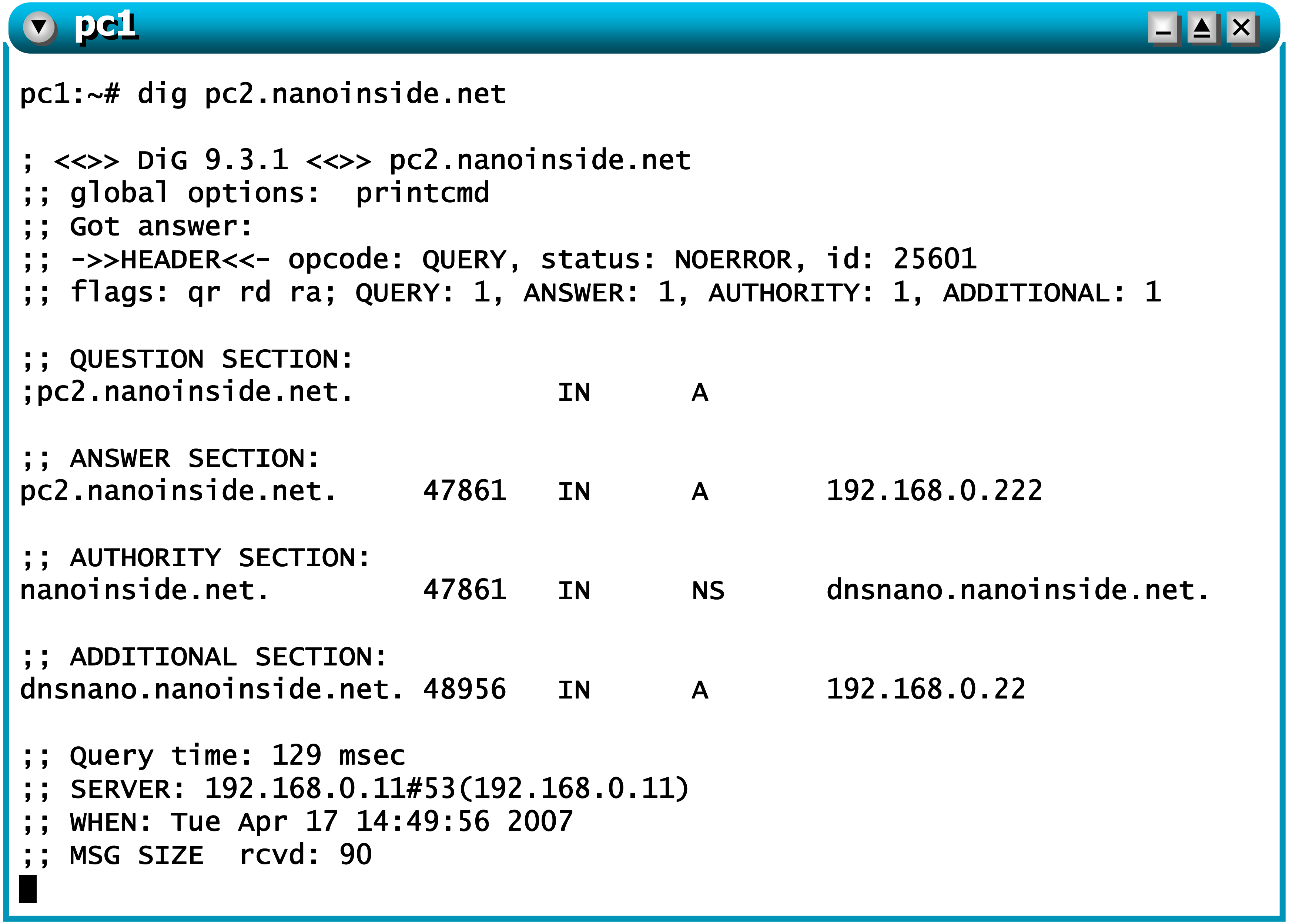
step 7 – advanced queries

- resource records can be searched by using **`dig`**
  - highly customizable queries
  - detailed responses



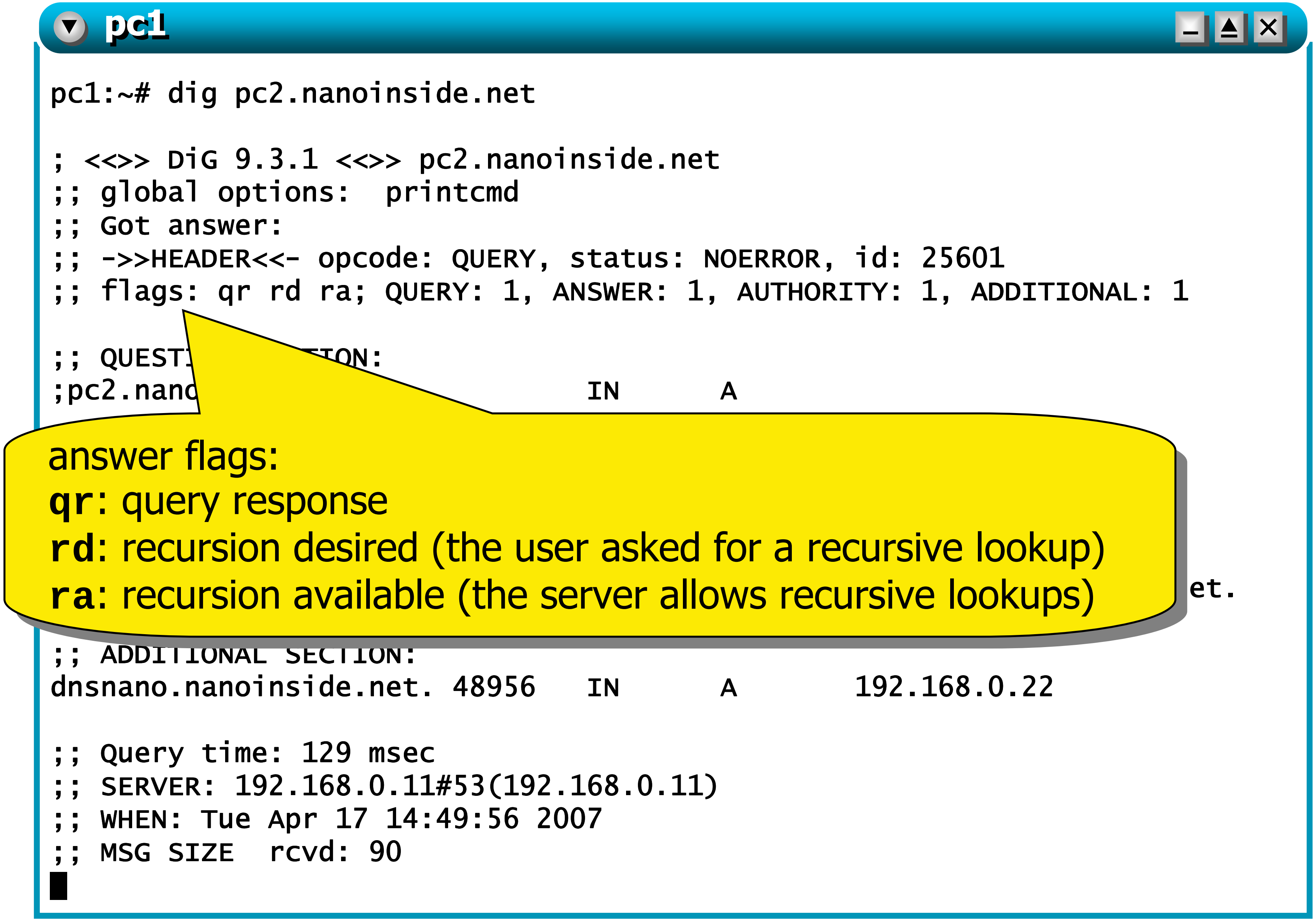
```
pc1:~# dig pc2.nanoinside.net █
```

step 7 – advanced queries



```
pc1:~# dig pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 25601
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; QUESTION SECTION:
;pc2.nanoinside.net.            IN      A

;; ANSWER SECTION:
pc2.nanoinside.net.     47861   IN      A       192.168.0.222

;; AUTHORITY SECTION:
nanoinside.net.         47861   IN      NS      dnsnano.nanoinside.net.

;; ADDITIONAL SECTION:
dnsnano.nanoinside.net. 48956   IN      A       192.168.0.22

;; Query time: 129 msec
;; SERVER: 192.168.0.11#53(192.168.0.11)
;; WHEN: Tue Apr 17 14:49:56 2007
;; MSG SIZE  rcvd: 90
```

step 7 – advanced queries



```
pc1:~# dig pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 25601
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; QUESTI          TION:
;pc2.nano                         IN      A

                                                                      et.
;; ADDITIONAL SECTION:
dnsnano.nanoinside.net. 48956   IN      A       192.168.0.22

;; Query time: 129 msec
;; SERVER: 192.168.0.11#53(192.168.0.11)
;; WHEN: Tue Apr 17 14:49:56 2007
;; MSG SIZE  rcvd: 90
```

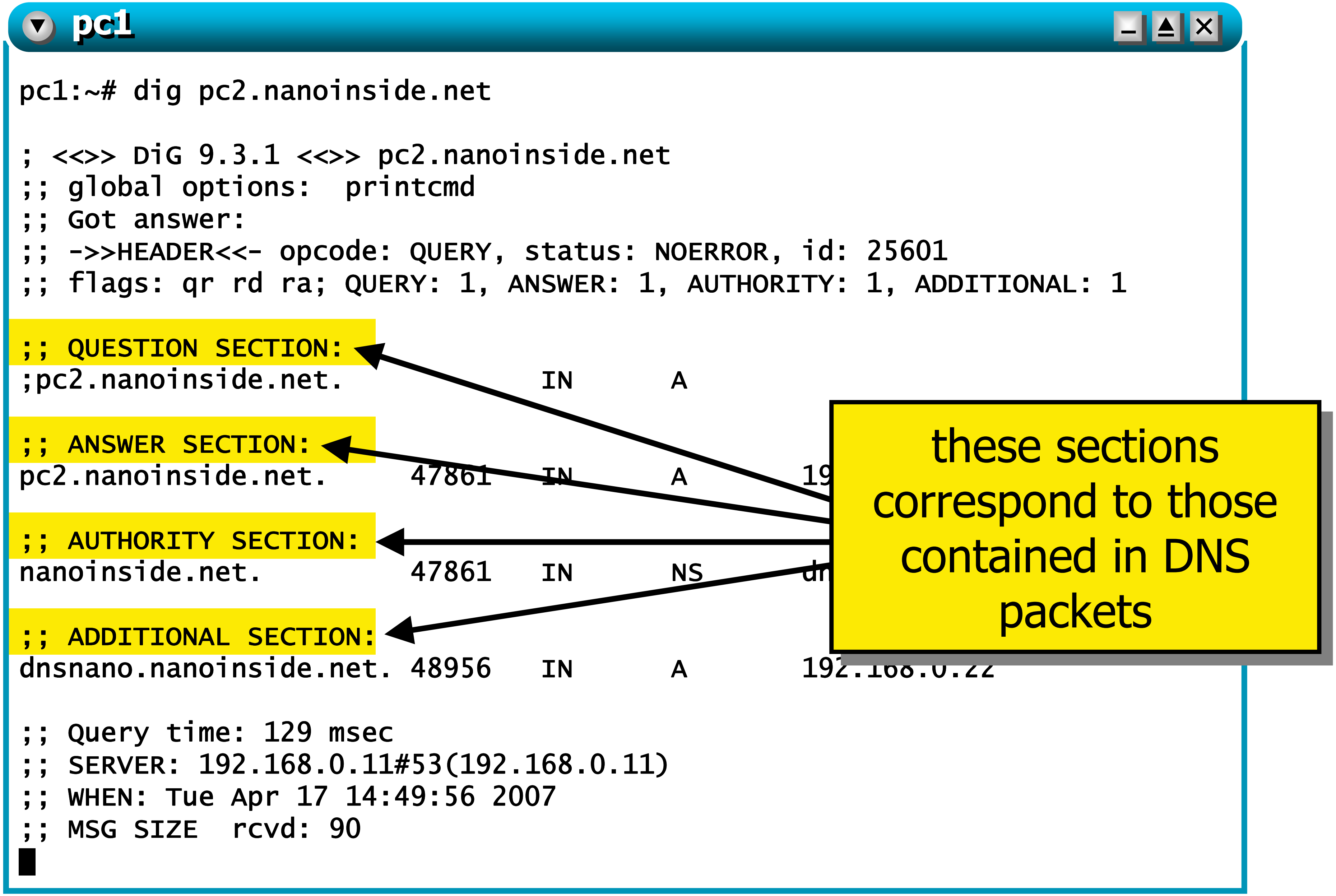
answer flags:

`qr`: query response

`rd`: recursion desired (the user asked for a recursive lookup)

`ra`: recursion available (the server allows recursive lookups)

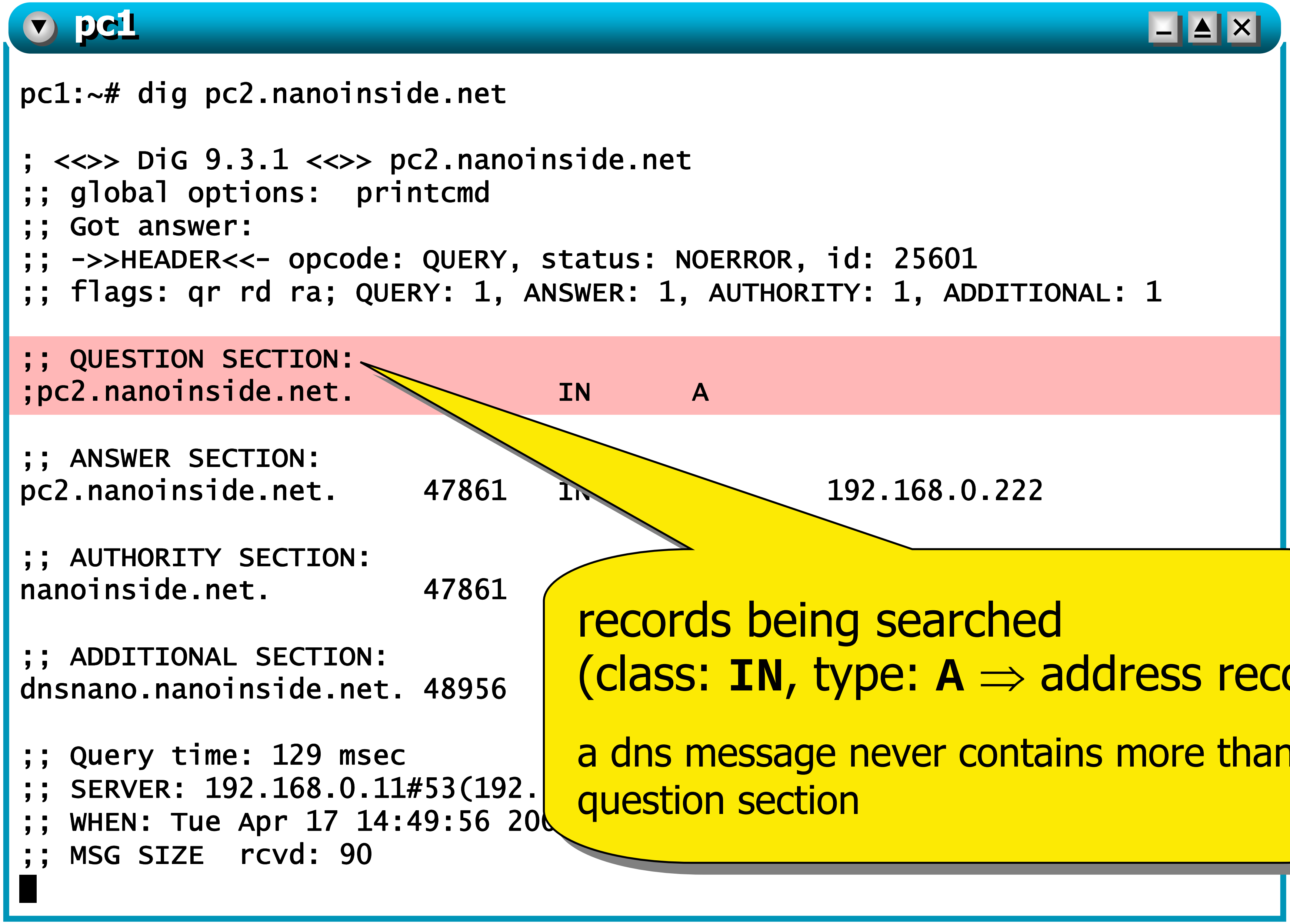
step 7 – advanced queries



```
pc1:~# dig pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 25601
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; QUESTION SECTION:
;pc2.nanoinside.net.            IN      A

;; ANSWER SECTION:
pc2.nanoinside.net.     47861   IN      A       19

;; AUTHORITY SECTION:
nanoinside.net.         47861   IN      NS      dn

;; ADDITIONAL SECTION:
dnsnano.nanoinside.net. 48956   IN      A       192.168.0.22

;; Query time: 129 msec
;; SERVER: 192.168.0.11#53(192.168.0.11)
;; WHEN: Tue Apr 17 14:49:56 2007
;; MSG SIZE  rcvd: 90
█
```

these sections correspond to those contained in DNS packets

step 7 – advanced queries

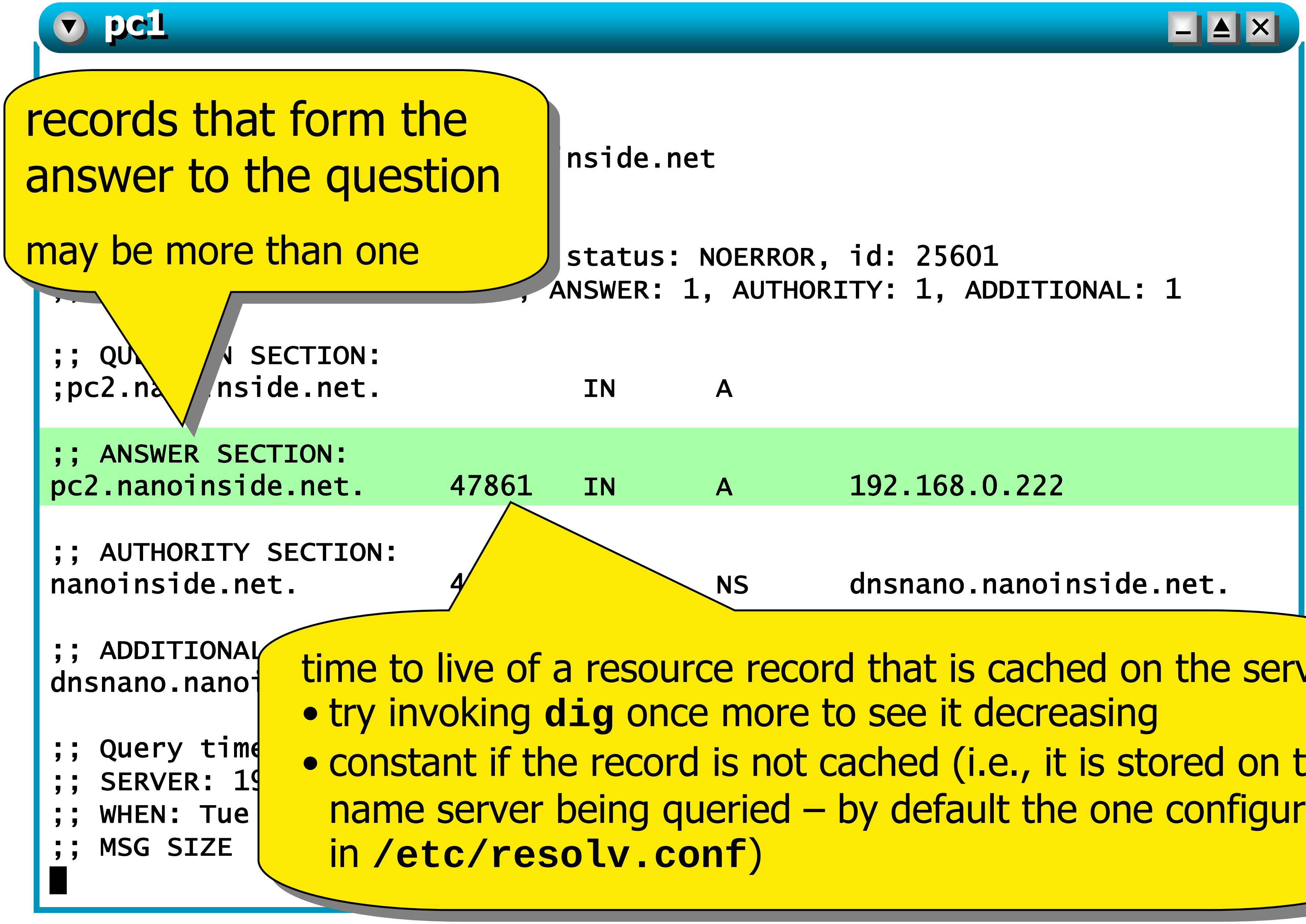


```
pc1:~# dig pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 25601
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; QUESTION SECTION:
;pc2.nanoinside.net.            IN      A

;; ANSWER SECTION:
pc2.nanoinside.net.     47861   IN              192.168.0.222

;; AUTHORITY SECTION:
nanoinside.net.         47861

;; ADDITIONAL SECTION:
dnsnano.nanoinside.net. 48956

;; Query time: 129 msec
;; SERVER: 192.168.0.11#53(192.
;; WHEN: Tue Apr 17 14:49:56 20
;; MSG SIZE  rcvd: 90
█
```

records being searched
(class: **IN**, type: **A** ⇒ address records)

a dns message never contains more than one question section

step 7 – advanced queries



```
                      nside.net

                      status: NOERROR, id: 25601
;;                  , ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; QU        N SECTION:
;pc2.na     nside.net.            IN      A

;; ANSWER SECTION:
pc2.nanoinside.net.     47861   IN      A       192.168.0.222

;; AUTHORITY SECTION:
nanoinside.net.         4               NS      dnsnano.nanoinside.net.

;; ADDITIONAL
dnsnano.nano

;; Query time
;; SERVER: 19
;; WHEN: Tue
;; MSG SIZE
```

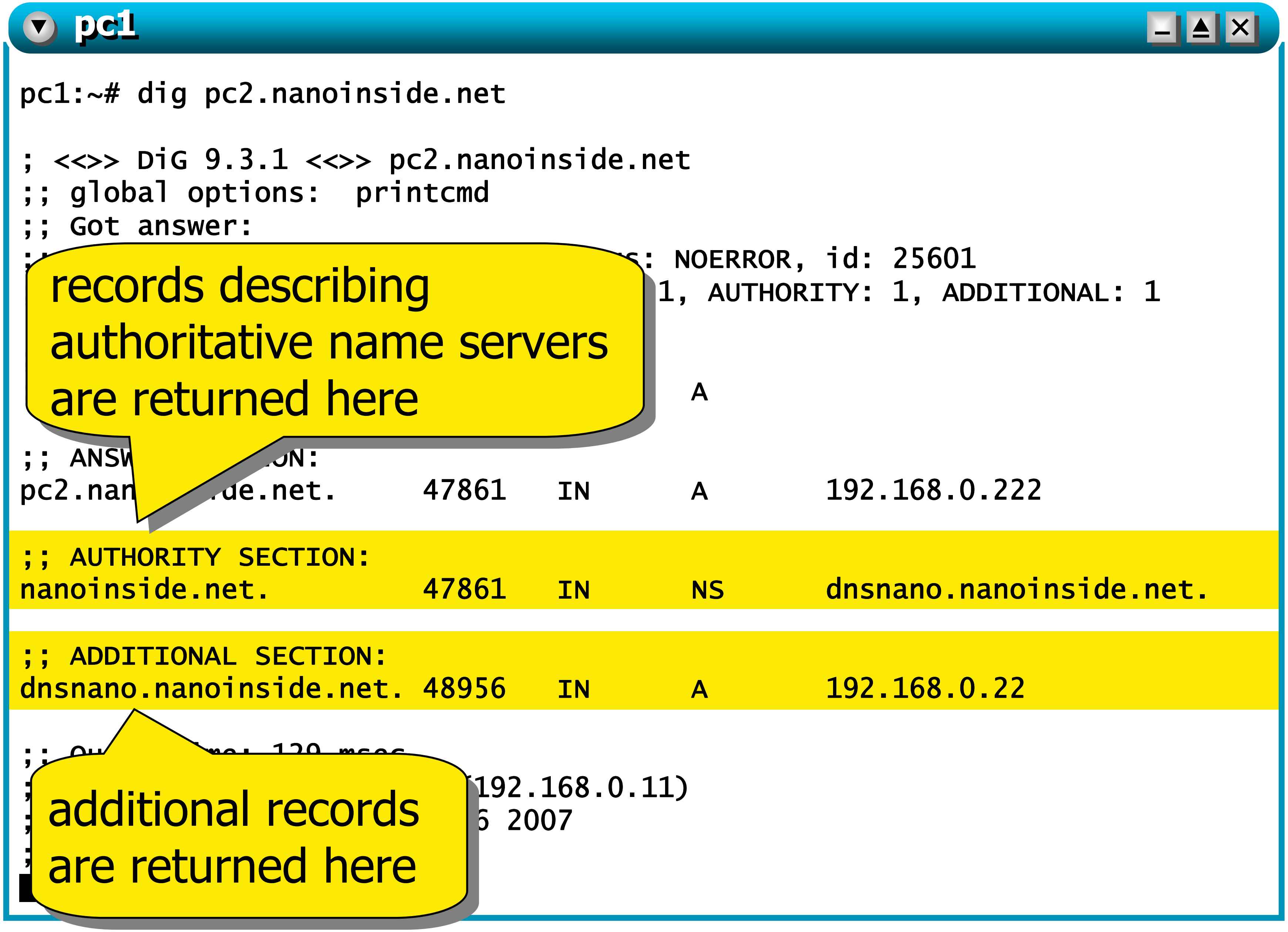
records that form the answer to the question

may be more than one

time to live of a resource record that is cached on the server

- try invoking **`dig`** once more to see it decreasing
- constant if the record is not cached (i.e., it is stored on the name server being queried – by default the one configured in **`/etc/resolv.conf`**)

step 7 – advanced queries

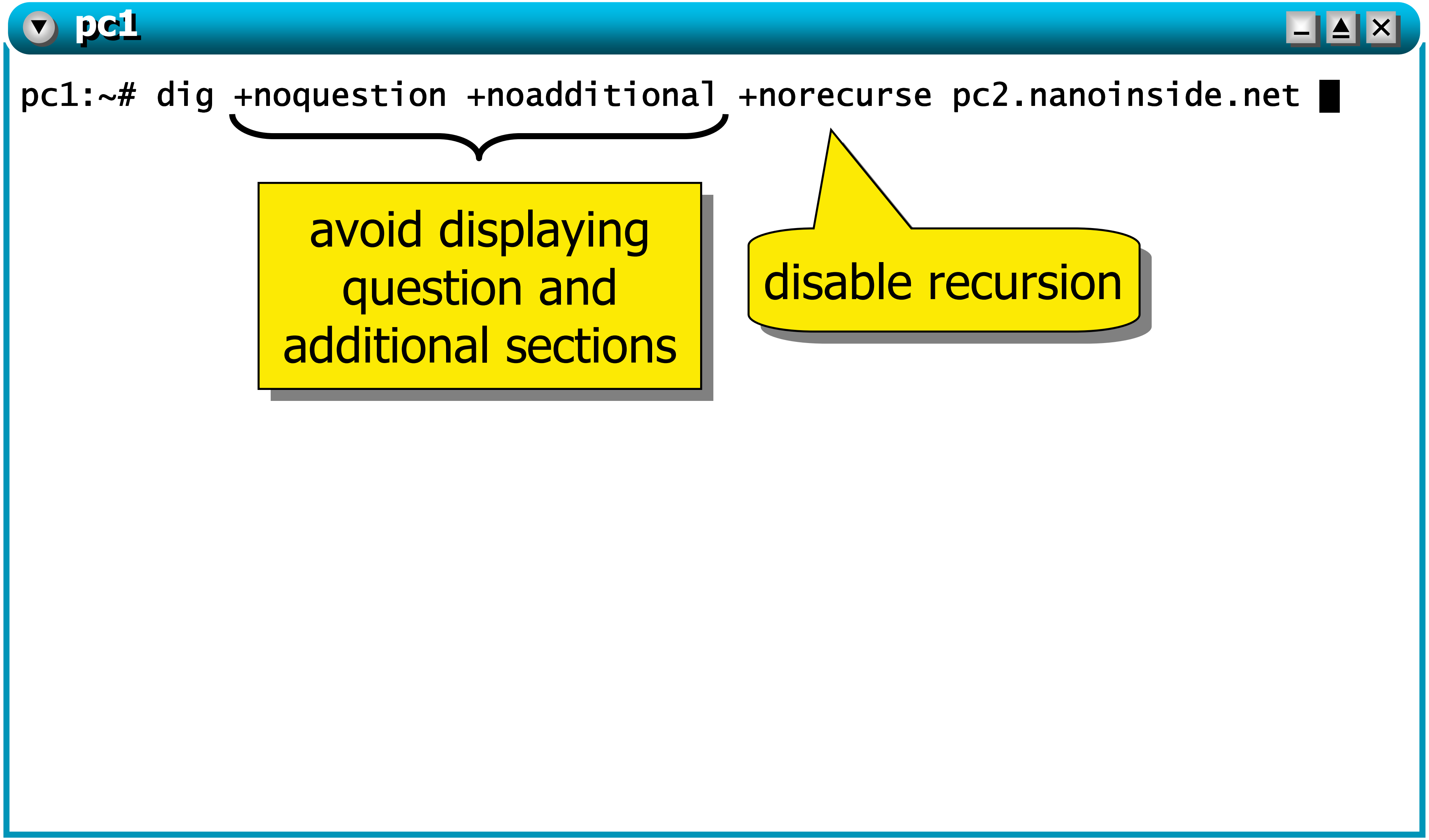


```
pc1:~# dig pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;;                                     s: NOERROR, id: 25601
;                                     1, AUTHORITY: 1, ADDITIONAL: 1



                                              A


;; ANSW       ON:
pc2.nan      de.net.        47861   IN      A       192.168.0.222

;; AUTHORITY SECTION:
nanoinside.net.             47861   IN      NS      dnsnano.nanoinside.net.

;; ADDITIONAL SECTION:
dnsnano.nanoinside.net. 48956   IN      A       192.168.0.22

;; Qu      me: 129 msec
;                          (192.168.0.11)
;                         6 2007
;
;
```

records describing authoritative name servers are returned here

additional records are returned here

step 8 – an iterative query

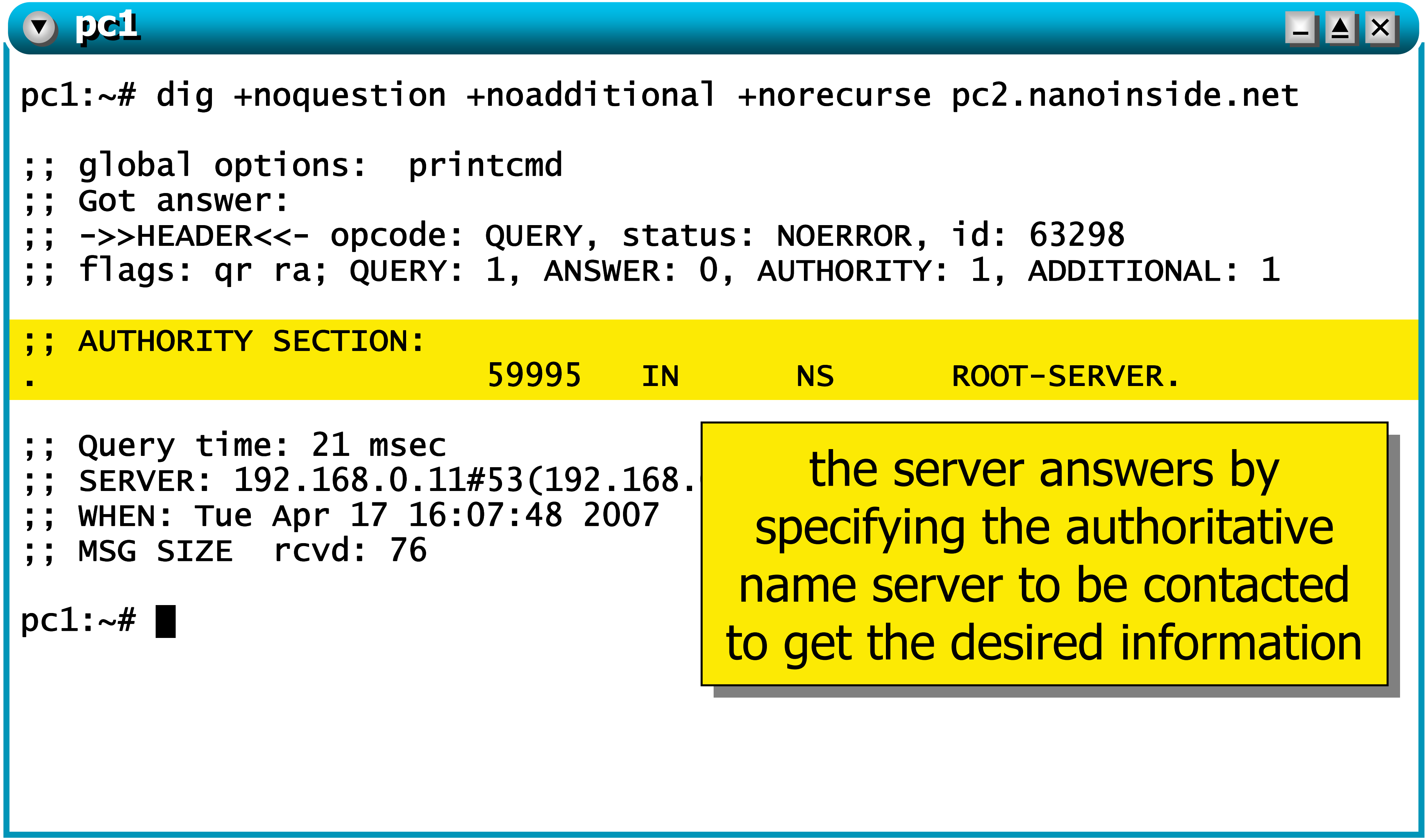


```
pc1:~# dig +noquestion +noadditional +norecurse pc2.nanoinside.net
```

avoid displaying question and additional sections

disable recursion

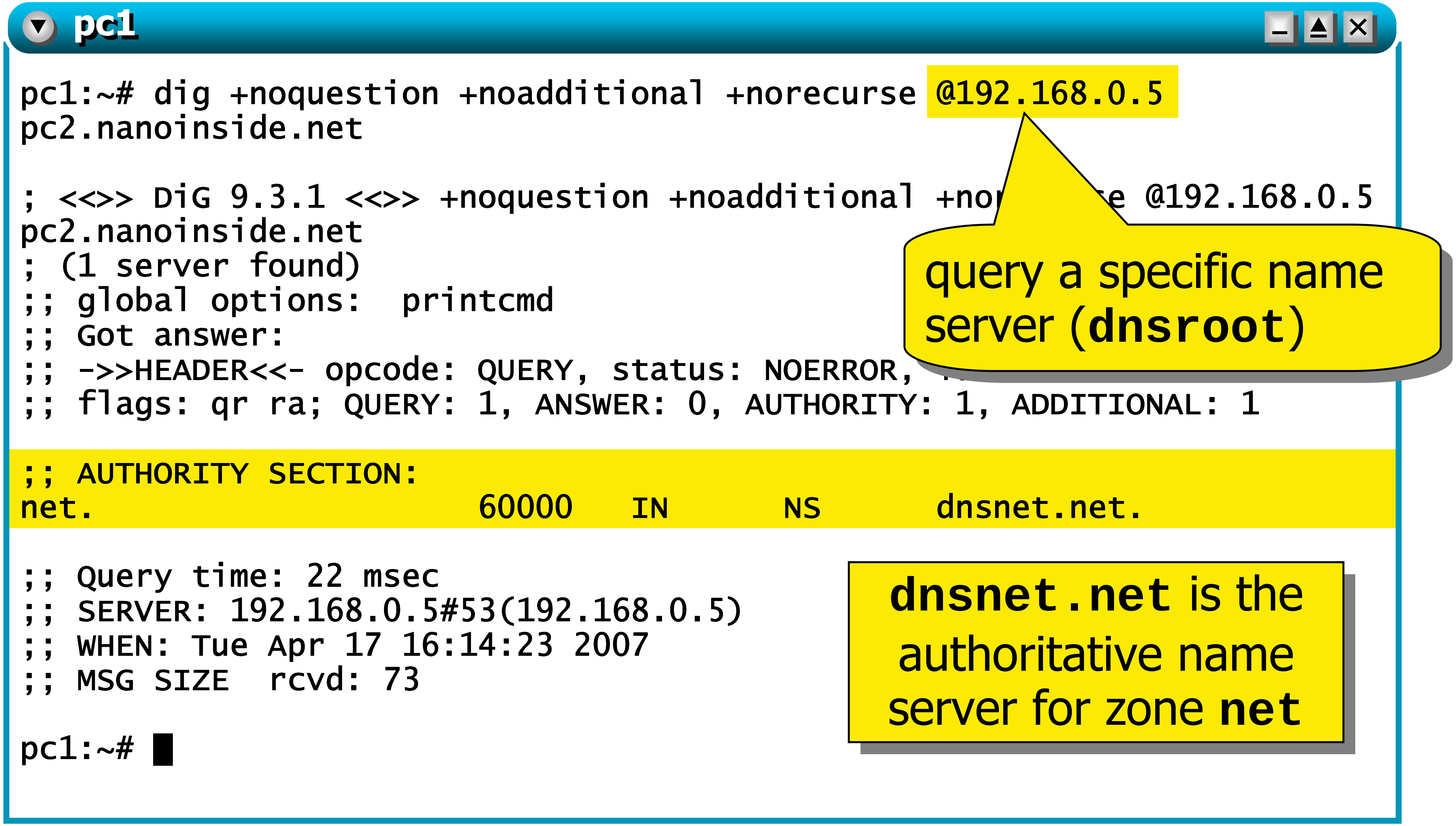
step 8 – an iterative query



```
pc1:~# dig +noquestion +noadditional +norecurse pc2.nanoinside.net

;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 63298
;; flags: qr ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; AUTHORITY SECTION:
.                       59995   IN      NS      ROOT-SERVER.

;; Query time: 21 msec
;; SERVER: 192.168.0.11#53(192.168.
;; WHEN: Tue Apr 17 16:07:48 2007
;; MSG SIZE  rcvd: 76

pc1:~# █
```

the server answers by specifying the authoritative name server to be contacted to get the desired information

step 8 – an iterative query

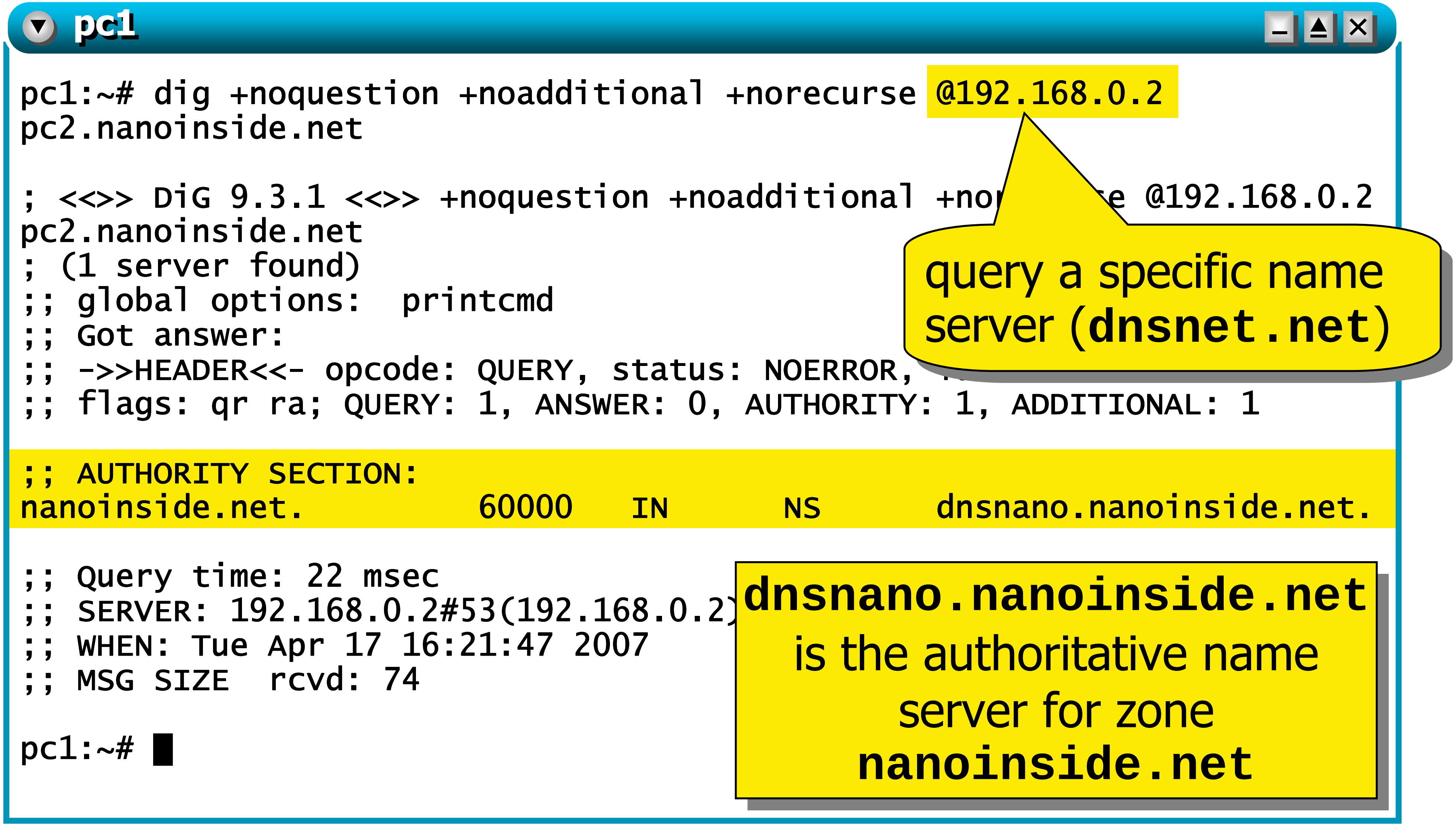


```
pc1:~# dig +noquestion +noadditional +norecurse @192.168.0.5
pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> +noquestion +noadditional +nor      e @192.168.0.5
pc2.nanoinside.net
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR,
;; flags: qr ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; AUTHORITY SECTION:
net.                        60000   IN      NS      dnsnet.net.

;; Query time: 22 msec
;; SERVER: 192.168.0.5#53(192.168.0.5)
;; WHEN: Tue Apr 17 16:14:23 2007
;; MSG SIZE  rcvd: 73

pc1:~# █
```

query a specific name server (**dnsroot**)

dnsnet.net is the authoritative name server for zone **net**

step 8 – an iterative query

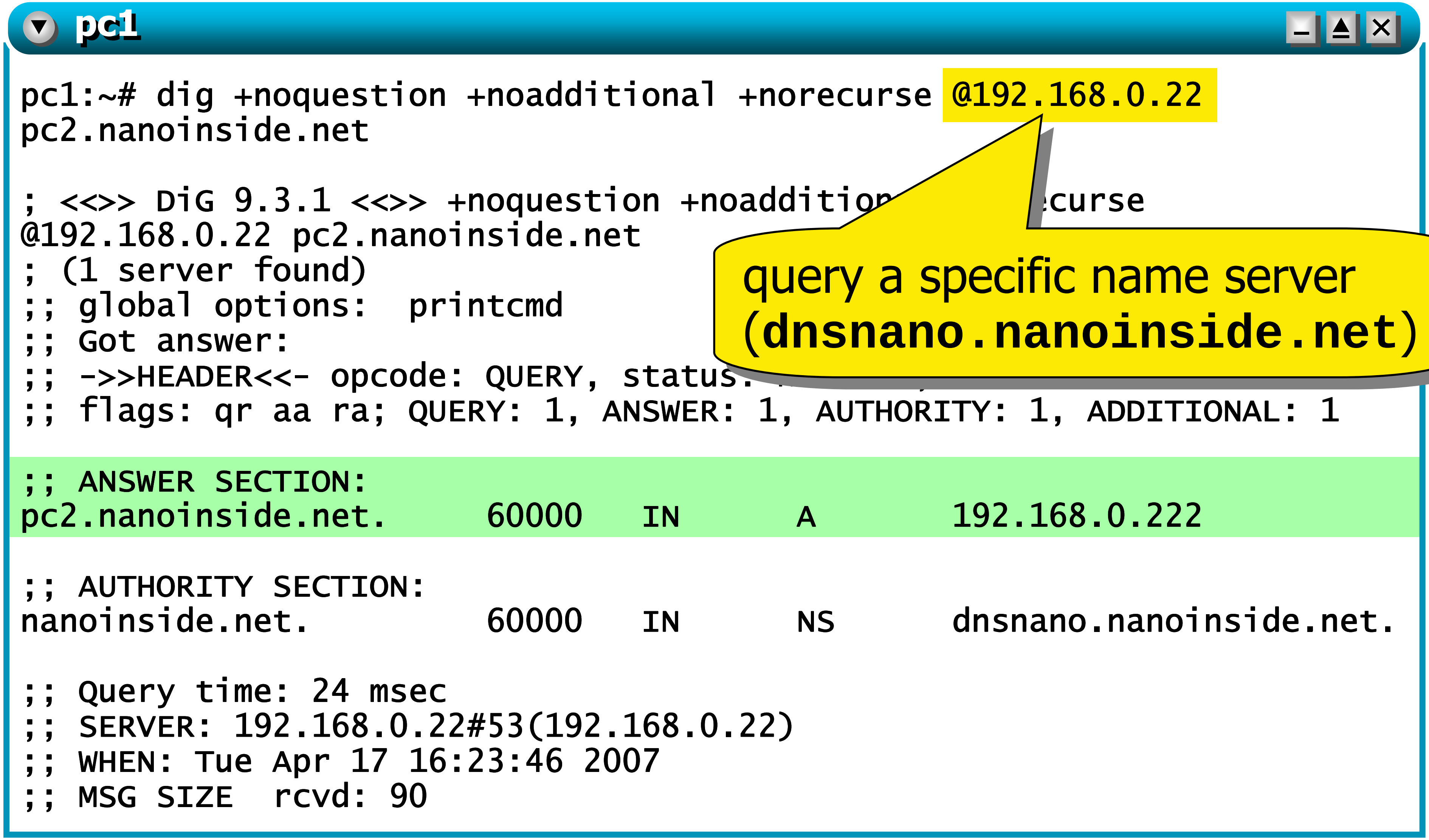


```
pc1:~# dig +noquestion +noadditional +norecurse @192.168.0.2
pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> +noquestion +noadditional +nor     e @192.168.0.2
pc2.nanoinside.net
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR,
;; flags: qr ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; AUTHORITY SECTION:
nanoinside.net.         60000   IN      NS      dnsnano.nanoinside.net.

;; Query time: 22 msec
;; SERVER: 192.168.0.2#53(192.168.0.2)
;; WHEN: Tue Apr 17 16:21:47 2007
;; MSG SIZE  rcvd: 74

pc1:~# █
```

query a specific name server (**dnsnet.net**)

dnsnano.nanoinside.net is the authoritative name server for zone **nanoinside.net**

step 8 – an iterative query

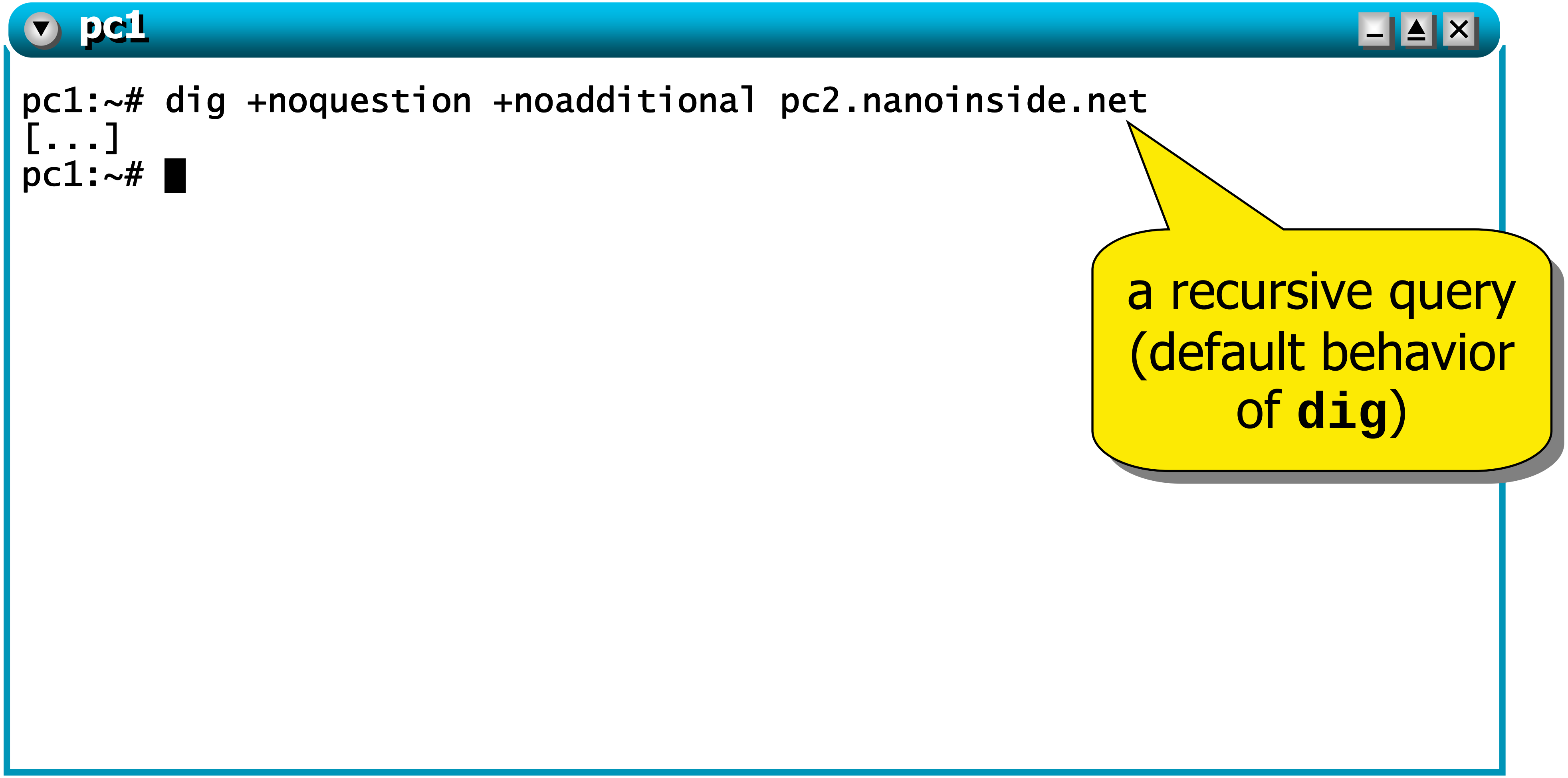


```
pc1:~# dig +noquestion +noadditional +norecurse @192.168.0.22
pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> +noquestion +noaddition            ecurse
@192.168.0.22 pc2.nanoinside.net
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status:
;; flags: qr aa ra; QUERY: 1, ANSWER: 1, AUTHORITY: 1, ADDITIONAL: 1

;; ANSWER SECTION:
pc2.nanoinside.net.     60000   IN      A       192.168.0.222

;; AUTHORITY SECTION:
nanoinside.net.         60000   IN      NS      dnsnano.nanoinside.net.

;; Query time: 24 msec
;; SERVER: 192.168.0.22#53(192.168.0.22)
;; WHEN: Tue Apr 17 16:23:46 2007
;; MSG SIZE  rcvd: 90
```

query a specific name server (**dnsnano.nanoinside.net**)

step 8 – an iterative query

- just to confirm that name servers cache information during recursive queries...

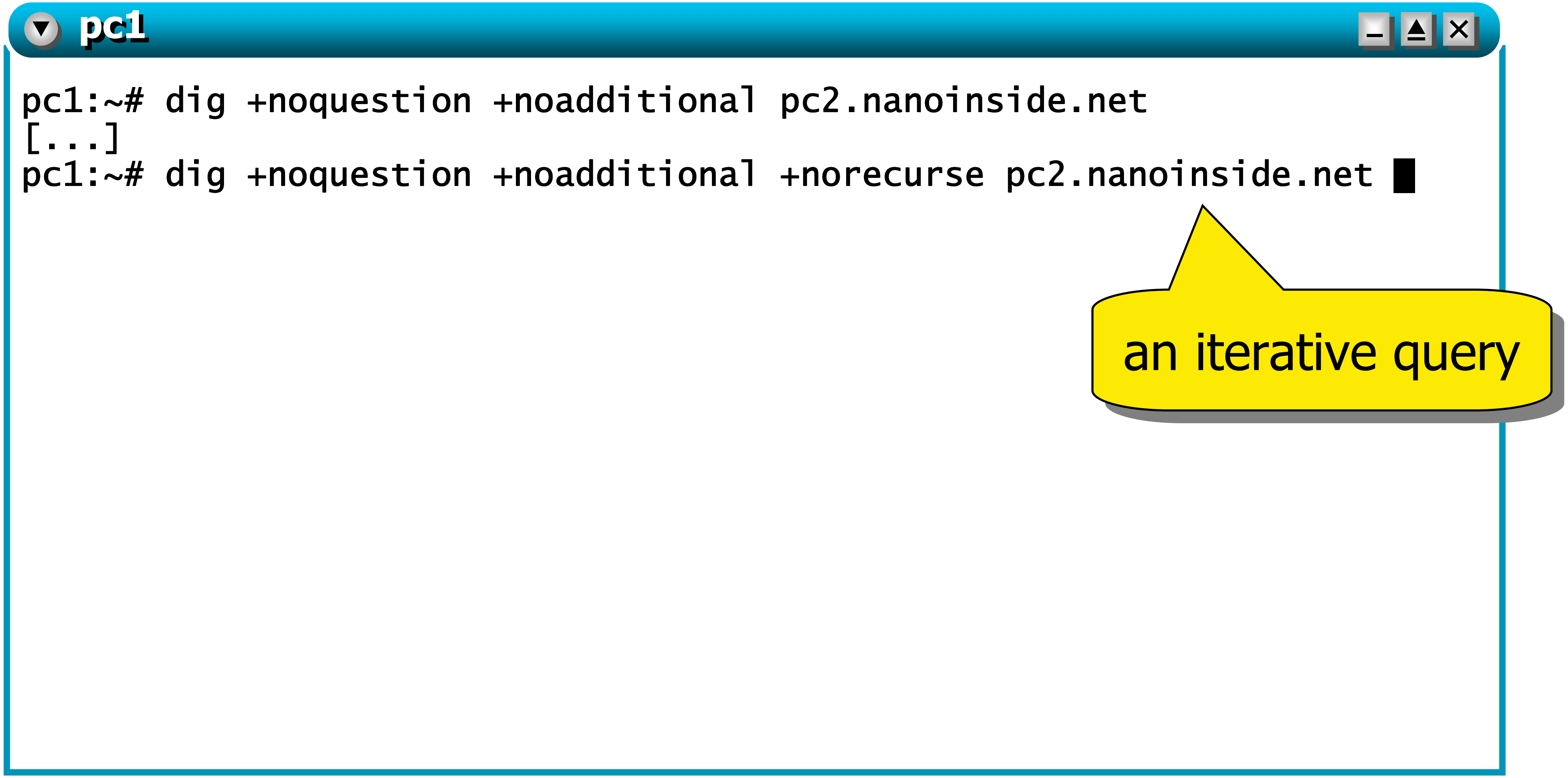


```
pc1:~# dig +noquestion +noadditional pc2.nanoinside.net
[...]
pc1:~# █
```

a recursive query (default behavior of **`dig`**)

step 8 – an iterative query

- just to confirm that name servers cache information during recursive queries...

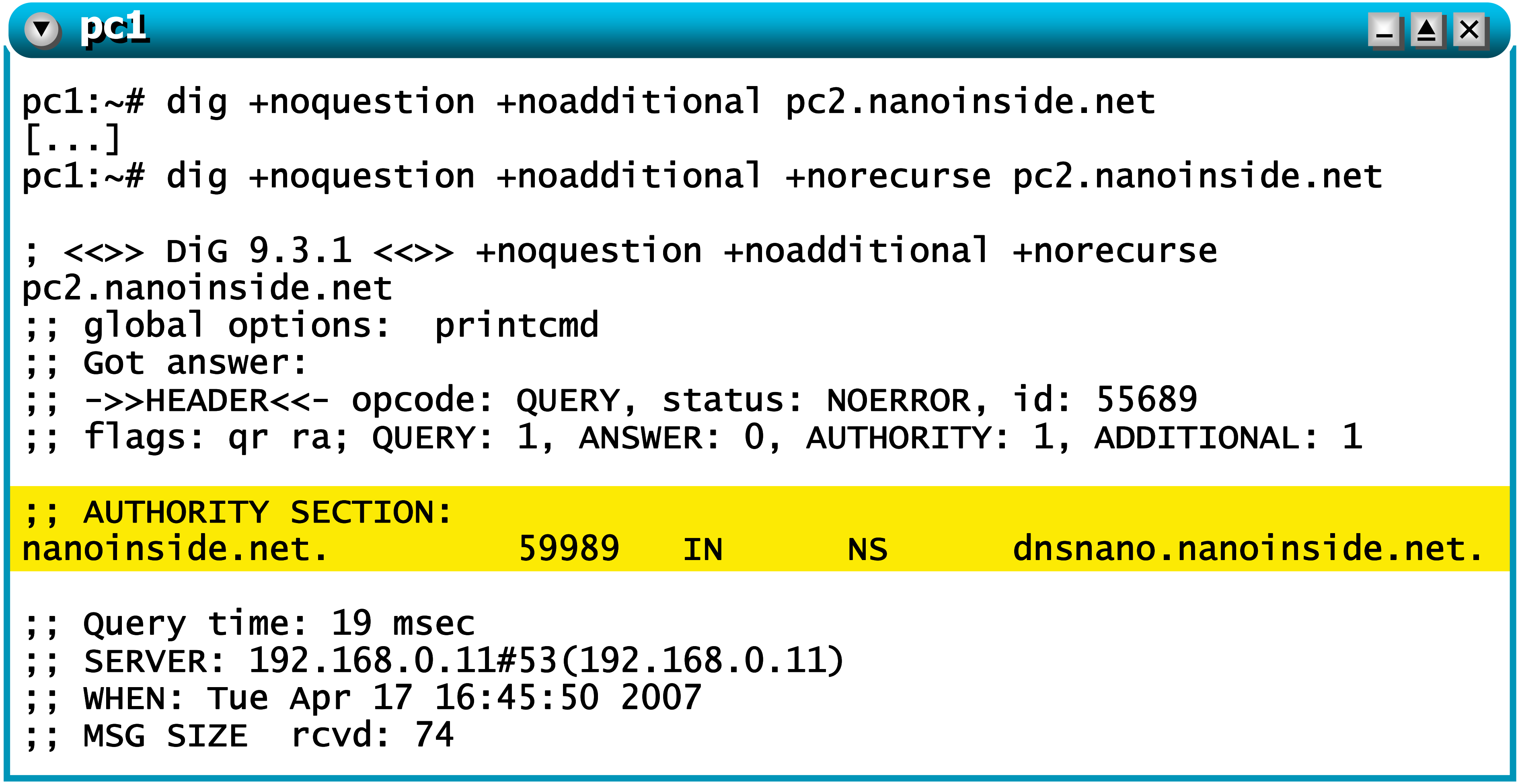


```
pc1:~# dig +noquestion +noadditional pc2.nanoinside.net
[...]
pc1:~# dig +noquestion +noadditional +norecurse pc2.nanoinside.net
```

an iterative query

step 8 – an iterative query

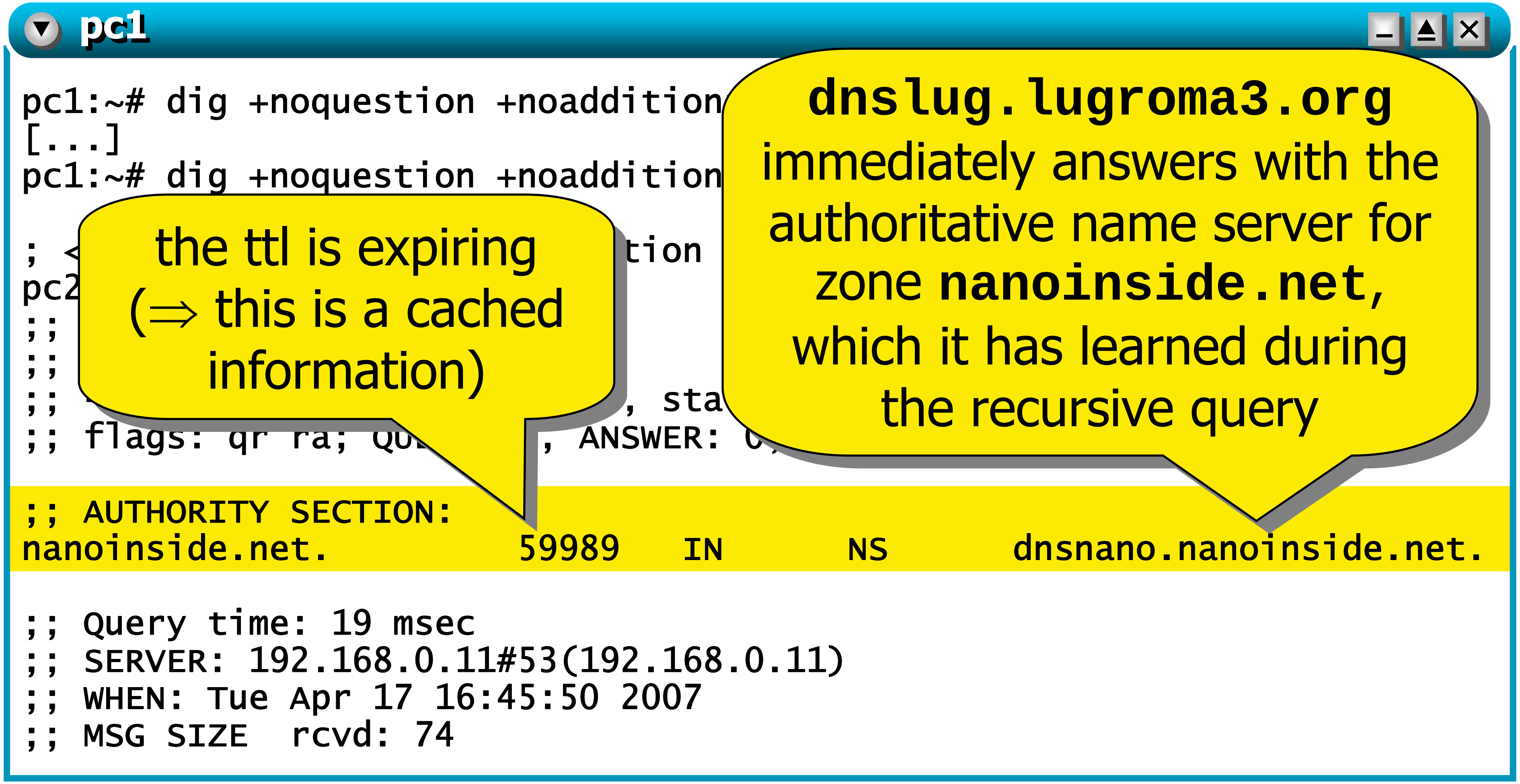
- just to confirm that name servers cache information during recursive queries...



```
pc1:~# dig +noquestion +noadditional pc2.nanoinside.net
[...]
pc1:~# dig +noquestion +noadditional +norecurse pc2.nanoinside.net

; <<>> DiG 9.3.1 <<>> +noquestion +noadditional +norecurse
pc2.nanoinside.net
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 55689
;; flags: qr ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; AUTHORITY SECTION:
nanoinside.net.         59989   IN      NS      dnsnano.nanoinside.net.

;; Query time: 19 msec
;; SERVER: 192.168.0.11#53(192.168.0.11)
;; WHEN: Tue Apr 17 16:45:50 2007
;; MSG SIZE  rcvd: 74
```

step 8 – an iterative query

- just to confirm that name servers cache information during recursive queries...



```
pc1:~# dig +noquestion +noaddition
[...]
pc1:~# dig +noquestion +noaddition

; <                                tion
pc2
;;
;;
;;                                , sta
;; flags: qr ra; QU               , ANSWER: 0,

;; AUTHORITY SECTION:
nanoinside.net.          59989   IN      NS      dnsnano.nanoinside.net.

;; Query time: 19 msec
;; SERVER: 192.168.0.11#53(192.168.0.11)
;; WHEN: Tue Apr 17 16:45:50 2007
;; MSG SIZE  rcvd: 74
```

the ttl is expiring (⇒ this is a cached information)

dnslug.lugroma3.org immediately answers with the authoritative name server for zone **nanoinside.net**, which it has learned during the recursive query