DNS
It's not short for Domain Name Screwups!

David Malone and Niall Murphy

21st February 2006

Names vs. Addresses

- Computers like addresses eg. 134.226.81.11.
- People prefer names salmon.maths.tcd.ie.
- Need a way to translate.
- walton.maths.tcd.ie close to salmon.maths.tcd.ie.
- 134.226.81.11 close to 134.226.81.12.
- DNS stores other data too.

DNS

The Domain Name System is a hierarchical distributed database.

- Hierarchy: www.google.com under google.com under com under '.'
- Distributed: responsibility delegated to 'owners'.
- Database: allows requesting types of information about domains.
- Contains information about *itself*.
- Is fundamental to Internet usage (by most folk).

Based on the notion of root servers.

Example query

Who is www.google.com? $\Rightarrow$ root server

$\Leftarrow$ com servers are ...

Who is www.google.com? $\Rightarrow$ com server

$\Leftarrow$ google.com servers are ...

Who is www.google.com? $\Rightarrow$ google server

$\Leftarrow$ www.google.com is ...

Example query types

For each domain, one can ask for:

A/AAAA IPv4/IPv6 Addresses.

TXT Comments.

NS Name of name server.

MX Name of mail server.

CNAME For real name.

PTR Name corresponding to address.

Reverse query

Want to translate 134.226.81.11 to it's name?

Look up 11.81.226.134.in-addr.arpa.

PTR for 11.81.226.134.in-addr.arpa? ⇒ root server

⇐ 134.in-addr.arpa servers

PTR for 11.81.226.134.in-addr.arpa? ⇒ 134.in-addr.arpa server

⇐ 226.134.in-addr.arpa servers

PTR for 11.81.226.134.in-addr.arpa? ⇒ 226.134.in-addr.arpa server

⇐ 81.226.134.in-addr.arpa server

PTR for 11.81.226.134.in-addr.arpa? ⇒ 81.226.134.in-addr.arpa server

⇐ salmon.maths.tcd.ie

Players in DNS game

Clients Make simple queries like 'MX for hotmail.com'.

Recursive Servers Answer simple queries by querying root and following delegation.

Authoritative Servers Know the answers (and delegations) for particular collection of names.

Two types of authoritative: masters (primary) and slaves (secondary). Slaves copy data from masters using zone transfer.

Less common players

Forwarders Sits between clients and recursive servers.

Stealth Secondaries Unadvertised secondaries, maybe on recursive servers.

Stealth Masters Sometimes you don't want to expose the real master.

Some of this is to do with caching.

Zones

- Chunk of DNS tree called zone.
- Headed by SOA record.
- Often represented by single file.
- Standard format — implied details.

```
$TTL    86400

@       IN      SOA     ns.maths.tcd.ie. hostmaster.maths.tcd.ie. (
                        ;       YYYYMMDDSS
                                2006020700 ; Serial
                                7200    ; Refresh
                                7200    ; Retry
                                604800  ; Expire
                                86400 ) ; (negative) TTL
        IN      NS      ns.maths.tcd.ie.
        IN      NS      ns1.tcd.ie.
        IN      NS      sec01.ns.esat.net.
        IN      LOC     53 20 34.9 N 6 15 0.5 W 30m 30m
        3600 IN MX      100     salmon
        3600 IN MX      300     kac.cnri.dit.ie.

ns      IN      A       134.226.81.11
        IN      AAAA    2001:770:10:300::86e2:510b

www     IN      CNAME   salmon

salmon  IN      A       134.226.81.11
        IN      HINFO   PC/Pentium              FreeBSD/4.2
        IN      MX      100     salmon
        IN      AAAA    2001:770:10:300::86e2:510b
```

Common zone mistakes

- Forgot to increment serial number (for BIND).
- Missing trailing '.'.
- CNAME and other data.
- CNAME/MX/NS to CNAME.
- _ in *hostname*.
- Uncontactable contact/MNAME.
- ';' is the comment character!
- Out of zone data.

Hooking into BIND

```
zone "maths.tcd.ie" {
        type master;
        file "p-i/maths.tcd.ie";
        also-notify {
                134.226.81.3; 134.226.81.8; 134.226.81.9;
                134.226.81.10; 134.226.81.12; 134.226.81.13;
                134.226.81.14; 134.226.81.15; 134.226.81.16;
                134.226.81.17; 134.226.81.18; 134.226.81.19;
                134.226.81.20; 134.226.81.21; };
        allow-transfer { any; };
};
```

Secondaries

```
zone "maths.tcd.ie" {
        type slave;
        file "s/maths.tcd.ie";
        masters { 134.226.81.11; };
};
```

Hooking into the tree

- Zone above needs to direct people to us.
- They need to duplicate NS records.
- They may need *glue*.

tcd zone:

```
maths                   NS      ns.maths
                        NS      ns1
                        NS      sec01.ns.esat.net.
ns.maths                A       134.226.81.11
                        AAAA    2001:770:10:300::86e2:510b
```

ie zone:

```
tcd.ie.                 172800  IN      NS      auth-ns1.ucd.ie.
tcd.ie.                 172800  IN      NS      ns.tcd.ie.
tcd.ie.                 172800  IN      NS      ns.maths.tcd.ie.
tcd.ie.                 172800  IN      NS      ns1.tcd.ie.
tcd.ie.                 172800  IN      NS      ns2.tcd.ie.
tcd.ie.                 172800  IN      NS      ns-sec.ripe.net.

ns.tcd.ie.              172800  IN      A       134.226.1.24
ns.maths.tcd.ie.        172800  IN      A       134.226.81.11
ns.maths.tcd.ie.        172800  IN      AAAA    2001:770:10:300::86e2:510b
ns1.tcd.ie.             172800  IN      A       134.226.1.114
ns2.tcd.ie.             172800  IN      A       134.226.1.28
auth-ns1.ucd.ie.        172800  IN      A       137.43.132.53
```

Delegation Mistakes

- Change NS/glue ⇒ upstream(s) update zone.
- Change master ⇒ secondaries update named.conf.
- Secondaries do AXFR. Don't ACL/firewall.
- Inconsistent answers from different machines/lame delegation.
- Inconsistent serials (with caveats).
- Cyclic dependencies (dependency 'footprint').
- Forgotten glue.

Reverse Zones

- Like forward, but ...
- Upstream for address space, not name space.
- Less problems with glue.
- Good for reverse and forward to be consistent.
- salmon.maths.tcd.ie.81.226.134.in-addr.arpa.
- Trickiest point: classless delegation.

Classic Reverse Zone

```
@       IN      SOA     ns.maths.tcd.ie. hostmaster.maths.tcd.ie. (
                        2006020700 86400 7200 604800 86400 )
        IN      NS      ns.maths.tcd.ie.
        IN      NS      ns1.tcd.ie.
        IN      NS      sec02.ns.esat.net.
1       IN      PTR     gw-81.maths.tcd.ie.
3       IN      PTR     lanczos.maths.tcd.ie.
8       IN      PTR     gosset.maths.tcd.ie.
9       IN      PTR     bell.maths.tcd.ie.
10      IN      PTR     walton.maths.tcd.ie.
11      IN      PTR     salmon.maths.tcd.ie.
```

named.conf

```
zone "81.226.134.in-addr.arpa" {
        type master;
        file "p-i/maths.rev-81";
        also-notify { 134.226.81.20; 134.226.81.21; };
        allow-transfer { any; };
};
```

RFC 2317 Parent Zone

```
@       IN      SOA     ns.maths.tcd.ie. hostmaster.maths.tcd.ie. (
                        2006020700 86400 7200 604800 86400 )
        IN      NS      ns.maths.tcd.ie.
        IN      NS      ns1.tcd.ie.
        IN      NS      sec02.ns.esat.net.

0-15              IN NS      ns1.customer.example.com.
0-15              IN NS      ns2.customer.example.com.
$GENERATE 1-15 $  IN    CNAME   $.0-15
```

RFC 2317 Child Zone

```
@       IN      SOA     ns.customer.example.com. hostmaster.customer.example.com. (
                        2006020700 86400 7200 604800 86400 )
1       IN      PTR     gw-81.customer.example.com.
3       IN      PTR     lanczos.customer.example.com.
8       IN      PTR     gosset.customer.example.com.
9       IN      PTR     bell.customer.example.com.
10      IN      PTR     walton.customer.example.com.
11      IN      PTR     salmon.customer.example.com.
```

DNS Best Practices

- In zone NSs (depends- -, efficency++);
- Usually good to spread resolvers around.
- Separate public authoritative and recursive service.
- Turn off unnecessary recursive service.
- Special zones: localhost. 0,127,255,10,168.192.in-addr.arpa. IPv6 equivalents (even if you don't use IPv6).
- Can use forward zones or stealth secondary.
- Update your root hints file.

```
From: Nick Hilliard <nick@IOL.IE>
Subject: Re: Call for DNS adjustment for IE domain
Date: Tue, 9 Jun 1998 09:36:23 +0100
To: IEDR-FORUM@LISTSERV.HEANET.IE

[...]
On another note, banba.ucd.ie appears to have recursion turned on.
As this server has no need for recursion (all it _should_ do is
authority service), can I suggest that recursion be turned off?

Nick



From: "Niall Richard Murphy (Sysadm)" <niallm@NETSOC.UCD.IE>
Subject: Re: Call for DNS adjustment for IE domain
Date: Tue, 9 Jun 1998 11:27:36 +0200
To: IEDR-FORUM@LISTSERV.HEANET.IE

[...]
Our opinion is, given
 * people are already using us for A & PTR resolution
 * the proportion of this 'real' resolution compared to just authority
   service is approximately 1/4
 * turning it off is effort better spent on other things
we should leave it the way it is.
```

From: Nick Hilliard <nick@IOL.IE>
Subject: Re: Call for DNS adjustment for IE domain
Date: Tue, 9 Jun 1998 12:21:30 +0100
To: IEDR-FORUM@LISTSERV.HEANET.IE

This is not a question of capacity -- ns.uu.net was quite happily running on a lowly sparc 2 (albeit with 128Mb of RAM) until about two years ago, and that was a pretty busy server. ns.eu.net was a sparc ELC for years. f.root-servers.net is a PC.

It is a question of security and applying the correct solution to the issue at hand.
1) RFC2010 is a useful yardstick in this case. [...]
2) Cache pollution problems surface regularly. [...]
3) Cache service and authority service are essentially different functions.
Turning off recursion is not necessary for servers like this. However, it is a very good idea.

Can I ask why people are using banba for general user resolution? Surely this is an IEDR machine limited to serving the needs of the IEDR only?

From: "Niall Richard Murphy (Sysadm)" <niallm@NETSOC.UCD.IE>
Subject: Re: Call for DNS adjustment for IE domain
Date: Tue, 9 Jun 1998 13:53:45 +0200
To: IEDR-FORUM@LISTSERV.HEANET.IE

> >Turning off recursion is not necessary for servers like this. However, it
> >is a very good idea.
At the moment, the way the registration system works requires banba to do recursion. [...]

Sometimes people...

- Restrict zone transfers.
- Leave out PTR, HINFO, ...
- Turn off version.bind.
- Use wildcards.

Digging for answers

```
> dig ns 2.0.0.2.ip6.int @z.ip6.int
; (2 servers found)
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 35489
;; flags: qr aa rd; QUERY: 1, ANSWER: 3, AUTHORITY: 0, ADDITIONAL: 4

;; QUESTION SECTION:
;2.0.0.2.ip6.int.               IN      NS

;; ANSWER SECTION:
2.0.0.2.ip6.int.        86400   IN      NS      flag.ep.net.
2.0.0.2.ip6.int.        86400   IN      NS      z.ip6.int.
2.0.0.2.ip6.int.        86400   IN      NS      dot.ep.net.

;; ADDITIONAL SECTION:
z.ip6.int.              86400   IN      A       198.32.2.66
z.ip6.int.              86400   IN      AAAA    3ffe:0:1::c620:242
flag.ep.net.            81749   IN      A       198.32.4.13
flag.ep.net.            81749   IN      AAAA    3ffe:805::2d0:b7ff:fee8:c4d9

;; Query time: 277 msec
;; SERVER: 3ffe:0:1::c620:242#53(3ffe:0:1::c620:242)
;; WHEN: Mon Feb 20 21:20:44 2006
;; MSG SIZE  rcvd: 180
```

```
> dig ns 2.0.0.2.ip6.int @3ffe:805::2d0:b7ff:fee8:c4d9
; (1 server found)
;; global options:  printcmd
;; connection timed out; no servers could be reached

> dig version.bind chaos txt @f.root-servers.net
;; ANSWER SECTION:
version.bind.           0       CH      TXT     "9.3.1"

> dig +trace +all aaaa www.maths.tcd.ie
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 18920
;; flags: qr ra; QUERY: 1, ANSWER: 13, AUTHORITY: 0, ADDITIONAL: 8

;; QUESTION SECTION:
;.                              IN      NS

;; ANSWER SECTION:
.                       451698  IN      NS      H.ROOT-SERVERS.NET.
.                       451698  IN      NS      I.ROOT-SERVERS.NET.
.                       451698  IN      NS      J.ROOT-SERVERS.NET.
.                       451698  IN      NS      K.ROOT-SERVERS.NET.
.                       451698  IN      NS      L.ROOT-SERVERS.NET.
.                       451698  IN      NS      M.ROOT-SERVERS.NET.
.                       451698  IN      NS      A.ROOT-SERVERS.NET.
.                       451698  IN      NS      B.ROOT-SERVERS.NET.
.                       451698  IN      NS      C.ROOT-SERVERS.NET.
.                       451698  IN      NS      D.ROOT-SERVERS.NET.
```

```
Use 'dig +trace +all ns 108.120.193.in-addr.arpa' found:

;; AUTHORITY SECTION:
120.193.in-addr.arpa.   172800  IN      NS      ns.ripe.net.
120.193.in-addr.arpa.   172800  IN      NS      sec02.ns.esat.net.
120.193.in-addr.arpa.   172800  IN      NS      auth02.ns.esat.net.

> dig ns 108.120.193.in-addr.arpa @ns.ripe.net
;; AUTHORITY SECTION:
108.120.193.in-addr.arpa. 86400 IN      NS      ns.ireland.eu.net.
108.120.193.in-addr.arpa. 86400 IN      NS      class.dublin.iona.ie.
108.120.193.in-addr.arpa. 86400 IN      NS      ns.maths.tcd.ie.

> dig ns 108.120.193.in-addr.arpa @sec02.ns.esat.net
;; ANSWER SECTION:
108.120.193.in-addr.arpa. 86400 IN      NS      auth02.ns.esat.net.
108.120.193.in-addr.arpa. 86400 IN      NS      sec02.ns.esat.net.

> dig ns 108.120.193.in-addr.arpa @auth02.ns.esat.net
;; ANSWER SECTION:
108.120.193.in-addr.arpa. 86400 IN      NS      auth02.ns.esat.net.
108.120.193.in-addr.arpa. 86400 IN      NS      sec02.ns.esat.net.
```

Other DNS stuff

- BIND views.
- Dynamic DNS updates (DDNS).
- Incremental Zone Transfer (IXFR).
- Anycast servers.
- IDNS.
- AAAA query rate increases (contrast BIND/Vista).
- DNSSEC (as we're about to see).

DNSSEC

- One definition: a way to be more confident that the answer you receive to a question is the intended answer.
- (of course, not necessarily the *correct* answer...)
- Another definition: extremely protracted, awkwardly-executed and still broken process.

DNSSEC — How does it work?

- Uses in band public-key cryptography.
- Aim: secure intra-resolver/authoritative server transactions.
- Sign a zone with your private key, publish public key in DNS.
- Heirarchical trust model — authenticity of maths.tcd.ie key established by tcd.ie.
- RFCs 4033,4034,4035.

- New resource records.
- RRSIG: a signature over a set of resource records with same name, class and type (www.maths.tcd.ie IN A).
- DNSKEY: public key, required for RRSIG verification.
- DS: pointers used in the trust model - parent is authoritative for the DS of the child zone.
- NSEC: internal 'next' pointers help with authenticated *non-existence* of data. (AXFR blocks now less relevant.)

TSIG

- Uses out-of-band keys and hash functions.
- Aim: secure intra-server transactions.
- Sign with shared secret and common timestamp.
- Compare supplied and calculated hash.
- Protects against replay, interception, alteration.
- Not scalable to arbitrary resolve/auth pairings.
- Requires synchronised clocks!
- Deployable separately from the rest of DNSSEC.
- RFC 2845.

DNSSEC — software support

- BIND 9 is your best bet.
- But it's not quite there yet (example from niallm).

DNSSEC — needs margarine

- A signed zone bloats rapidly in size and loading time.
- Eg: .nl goes from 40Mb unsigned to 350Mb+ signed.
- Eg: Signing time: 1.5 *hours*
- Eg: Loading time: 15 minutes.
- Apparently .com goes to 10 Gb when you do this.
- Likely not to be a problem for you and me.

Comments

p6 SPF/ClamAV use of TXT records.

p7 One way to think of the types of servers is that your recursive servers serve your network, your authoritative servers "advertise" your domain to other networks.

p7 A crucial point about recursive servers is that they can cache results, and DNS provides for this via TTLs.

p8 Forwarders/hidden masters are often used to accommodate machines behind firewalls.

p8 In some cases, it may be a database that you hide rather than a master handing out AXFRs.

p9 Zone file format is actually covered in the RFCs.

p10 If using YYYYMMDDSS for the serial number, maybe put "YYYYmmDDss" above as a comment, so you can clearly see the boundaries.

p10 LOC records can be used by geographic traceroute.

p10 HINFO record (were/are) used by multinet telnet.

p11 _ are not permitted in any part of a hostname. There are lots of other illegal characters too.

p11 One common cause of out-of-zone data is a trailing dot where one was not intended.

p13 Why duplicate NS records in zone? Glue records are non-authoritative.

p15 DJB's tool "dnstrace" will show all paths through the

dns tree to a given domain starting from a given server. Like "dig +trace +all", but checks all paths.

p16 Name servers with names in in-addr.arpa do exist, to keep the name server in-zone.

p19 It is good to spread both authoritative servers around then Internet and recursive servers around your network. It means your domain continues to exist when you're off the network *and* people in your network can continue to work when your network is partitioned.

p19 How often to update your root hints file? Yearly seems like a reasonable choice.

p22 If zones contain personal information, then it may be wise to restrict zone transfers for data protection reasons.

p23 Wildcards records are actually part of the spec. May be possible to query for them. (Note, Wikipedia suggests that all DNS servers implement wildcards in different ways, none of which match RFC 1034!)

p26 Incremental zone transfer now available for non-dynamic zones in recent versions of BIND. It will build the .jnl file from your manual edits.

p26 International domain names pose some interesting problems in terms of politics and security (glyphs that look the same but are represented differently).

wrap up What are some good tools for checking DNS? Suggestions

dnswalk — `http://www.visi.com/~barr/dnswalk/`

dnscheck — `http://www.dnscheck.se/`

zonecheck — `http://www.zonecheck.fr/`

named-checkconf, named-checkzone — part of BIND.