DNS Session 4: Delegation and reverse DNS

Joe Abley
AfNOG 2007 workshop

How do you delegate a subdomain?

- In principle straightforward: just insert NS records for the subdomain, pointing at someone else's servers
- If you are being careful, you should first *check that those servers are authoritative for the subdomain*
  - *by using "dig +norec" on all the servers*
- *If the subdomain is managed badly, it reflects badly on you!*
  - *and you don't want to be fielding problem reports when the problem is somewhere else*

Zone file for "example.com"

```
$TTL 1d
@  1h  IN  SOA  ns1.example.net. brian.nsrc.org. (
             2004030300     ; Serial
             8h             ; Refresh
             1h             ; Retry
             4w             ; Expire
             1h )           ; Negative

       IN  NS  ns1.example.net.
       IN  NS  ns2.example.net.
       IN  NS  ns1.othernetwork.com.

; My own zone data
       IN  MX  10  mailhost.example.net.
www    IN  A   212.74.112.80

; A delegated subdomain
subdom IN  NS  ns1.othernet.net.
       IN  NS  ns2.othernet.net.
```

There is one problem here:

- NS records point to names, not IPs
- What if zone "example.com" is delegated to "ns.example.com"?
- Someone who is in the process of resolving (say) www.example.com first has to resolve ns.example.com
- But in order to resolve ns.example.com they must first resolve ns.example.com !!

In this case you need "glue"

- A "glue record" is an A record for the nameserver, held higher in the tree
- Example: consider the .com nameservers, and a delegation for example.com

```
; this is the com. zone

example          NS  ns.example.com.
                 NS  ns.othernet.net.

ns.example.com.  A   192.0.2.1     ; GLUE RECORD
```

Don't put in glue records except where necessary

- In the previous example, "ns.othernet.net" is not a subdomain of "example.com". Therefore no glue is needed.
- Out-of-date glue records are a big source of problems
  - e.g. after renumbering a nameserver
  - Results in intermittent problems, difficult to debug

Example where a glue record IS needed

```
; My own zone data
            IN  MX  10  mailhost.example.net.
www         IN  A   212.74.112.80

; A delegated subdomain
subdom      IN  NS  ns1.subdom          ; needs glue
            IN  NS  ns2.othernet.net.   ; doesn't
ns1.subdom  IN  A   192.0.2.4
```

Checking for glue records

- dig +norec ... *and repeat several times*
- Look for A records in the "Additional" section whose TTL does not count down

```
$ dig +norec @a.gtld-servers.net. www.as9105.net. a
 ...
;; flags: qr; QUERY: 1, ANSWER: 0, AUTHORITY: 2, ADDITIONAL: 1
;; QUERY SECTION:
;;      www.as9105.net, type = A, class = IN

;; AUTHORITY SECTION:
as9105.net.          172800  IN    NS        ns0.as9105.com.
as9105.net.          172800  IN    NS        ns0.tiscali.co.uk.

;; ADDITIONAL SECTION:
ns0.as9105.com.      172800  IN    A         212.139.129.130
```

Practical

- Delegating a subdomain

Loose ends: how to manage reverse DNS

- If you have at least a /24 of address space then your provider will arrange delegation to your nameservers
- e.g. your netblock is 196.222.0.0/24
- Set up zone 0.222.196.in-addr.arpa.
- If you have more than a /24, then each /24 will be a separate zone
- If you a lucky enough to have a /16 then it will be a single zone
  - 196.222.0.0/16 is 222.196.in-addr.arpa.

Example: 196.222.0/24

/etc/namedb/named.conf

```
zone "0.222.196.in-addr.arpa" {
   type master;
   file "master/196.222.0";
   allow-transfer { ... };
};
```

/etc/namedb/master/196.222.0

```
@   IN  SOA  ....
    IN  NS   ns0.example.com.
    IN  NS   ns0.othernetwork.com.

1   IN  PTR  router-e0.example.com.
2   IN  PTR  ns0.example.com.
3   IN  PTR  mailhost.example.com.
4   IN  PTR  www.example.com.
; etc
```

How it works

- e.g. for 196.222.0.4, the remote host will lookup 4.0.222.196.in-addr.arpa. (PTR)
- The query follows the delegation tree as normal. If all is correct, it will reach your nameservers and you will reply
- Now you can see why the octets are reversed
  - The owner of a large netblock (e.g. 192/8) can delegate reverse DNS in chunks of /16. The owner of a /16 can delegate chunks of /24

There is nothing special about reverse DNS

- You still need master and slave(s)
- It won't work unless you get delegation from above
- However, DO make sure that if you have a PTR record for an IP address, that the hostname resolves back to the same IP address
  - Otherwise, many sites on the Internet will think you are spoofing reverse DNS and will refuse to let you connect

What if you have less than /24?

- Reverse DNS for the /24 has been delegated to your upstream provider
- Option 1: ask your provider to insert PTR records into their DNS servers
  - Problem: you have to ask them every time you want to make a change
- Option 2: follow the procedure in RFC 2317
  - Uses a trick with CNAME to redirect PTR requests for your IPs to your nameservers

e.g. you own 192.0.2.64/29

In the provider's 2.0.192.in-addr.arpa zone file

```
64     IN  CNAME  64.64/29.2.0.192.in-addr.arpa.
65     IN  CNAME  65.64/29.2.0.192.in-addr.arpa.
66     IN  CNAME  66.64/29.2.0.192.in-addr.arpa.
67     IN  CNAME  67.64/29.2.0.192.in-addr.arpa.
68     IN  CNAME  68.64/29.2.0.192.in-addr.arpa.
69     IN  CNAME  69.64/29.2.0.192.in-addr.arpa.
70     IN  CNAME  70.64/29.2.0.192.in-addr.arpa.
71     IN  CNAME  71.64/29.2.0.192.in-addr.arpa.
64/29  IN  NS   ns0.customer.com.
64/29  IN  NS   ns1.customer.com.
```

Set up zone **"64/29.2.0.192.in-addr.arpa"** on your nameservers

```
65     IN  PTR    www.customer.com.
66     IN  PTR    mailhost.customer.com.
; etc
```

DNS: Summary

- Distributed database of Resource Records
  - e.g. A, MX, PTR, ...
- Three roles: resolver, cache, authoritative
- Resolver statically configured with nearest caches
  - e.g. /etc/resolv.conf
- Caches are seeded with a list of root servers
  - zone type "hint", /etc/namedb/named.root
- Authoritative servers contain RRs for certain zones (part of the DNS tree)
  - replicated for resilience and load-sharing

DNS: Summary (cont)

- Root nameservers contain delegations (NS records) to gTLD or country-level servers (com, uk etc)
- These contain further delegations to subdomains
- Cache finally locates an authoritative server containing the RRs requested
- Errors in delegation or in configuration of authoritative servers result in no answer or inconsistent answers

Further reading

- "DNS and BIND" (O'Reilly)
- BIND 9 Administrator Reference Manual
  - /usr/share/doc/bind9/arm/Bv9ARM.html
- http://www.isc.org/sw/bind/
  - includes FAQ, security alerts
- RFC 1912, RFC 2182
  - http://www.rfc-editor.org/