The DNS Today
Are we Overloading the Saddlebags on an Old Horse?

Randy Bush <randy@psg.com>
IETF / San Diego 00.12.13
Double Entendre

What’s new in the DNS

Architectural Restraint

The DNS may be said to be the major fairly-well-distributed store of the internet. We are now looking at rather clever extensions for IPv6, security, etc. Will they work? Scale?
Delay, as the User Sees it

Connect: Looking up host: somewhere-on-the.net...
IPv6 Support (RFC 2874)

- AAAA (RFC 1886, deprecated)
- A6 (RFC 2874)
- Binary Labels (RFC 2637)
- DNAME (RFC 2672)
AAAA (RFC 1886) (deprecated)

foo AAAA 666:0:1:2:3:4:567:89ab
$ORIGIN X.EXAMPLE.

N A6 64 ::1234:5678:9ABC:DEF0 SN-1.IP6
SN-1.IP6 A6 48 0:0:0:1:: IP6
IP6 A6 48 0::0 CUST-X.IP6.A.NET.
IP6 A6 48 0::0 CUST-X.IP6.B.NET.
And elsewhere

$ORIGIN NET.

CUST-X.IP6.A A6 40 0:0:0011:: A.NET.IP6.D
A.NET.IP6.C A6 28 0:0001:CA00:: C.NET.ORG.
A.NET.IP6.D A6 28 0:0002:DA00:: D.NET.ORG.
B-NET.IP6.E A6 32 0:0:EB00:: E.NET.ORG.
C.NET.ORG. A6 0 2345:00C0::
D.NET.ORG. A6 0 2345:00D0::
E.NET.ORG. A6 0 2345:000E::

And if you think that's complex, you should see the example with glue RRs
Binary Labels (RFC 2673)

\[b11010000011101]\n\[o64072/14]\n\[xd074/14]\n\[208.116.0.0/14]\n
The following represents two consecutive Bit-String Labels which denote the same relative point in the DNS tree as any of the above single Bit-String Labels.
\[b11101\].\[o640\]
DNAME (RFC 2672)
(please use only for renumbering)

190.189.188.1 reverse 1.188.189.190.in-addr.arpa

$ORIGIN new-style.in-addr.arpa.
189.190 DNAME in-addr.example.net.
$ORIGIN in-addr.example.net.
188 DNAME in-addr.customer.example.
$ORIGIN in-addr.customer.example.
1 PTR www.customer.example.
2 PTR mailhub.customer.example.

Think about DNAME for binary with CIDR
DNSsec Security
(RFCs 2535 2931)

- Authenticates data, not servers
- KEY - public keys, private keys off-line
- SIG - signs RRs
- NXT - fills empty space
KEY

KEY 256 3 3 ( // flags, protocol, algorithm
CPZIUu8BpuNfNybf5Fobd7W26+rjxe9sBUtCzc0cJmim
hbFjwWUM5fbUmgJEcwK1e1D86PP+Yg0K/QceZb3Pstap
613uOZokBjCvFhosY9LLJaIrmYULvo6Q0mizsF79GZfV
bk10vDfOCZaA8ehURnTTWxz8LiikAXOouGNvDQBSX0tU
epISThM4pEuPW1mjGrL+ukU8BRk+PRb/dnVF4D8MkEvN
wnLz8Bc9t+dzKrZIBodVaKCV689hQh67uA0TDHG4fZrR
eweYifAwnGuWvqd9Tmt8zOrBEIx1qmGebfKYCa9ecVS3c
m799N8WhS9rtSlczhd3YJ15jrMoMXFdvohNJwNiMYtye
jjeeOK0Jfsiv9v6ITVty170gs+WkD4FDxJUKUVEpScfG
9+R2E05oE9SOHM62uE8sgG2PjGVEQ+BLQ/q05bEU2qxS
op30y1Gn2COeh1xy1J3f4M/ikDJHooQprjFtnd5C6rhH
HhFf3Tw7BILwXvhW41wkoCnZjjwY9So3jm/ws6Jek7/X
Vm48aptkfFD7 )
// SIG type, algorithm, label, ottl, expires, created, tag, signer

SIG SOA 3 2 60 20001008092222 (  
  20000908092222 59140 randy.se.  
  CAs/Br+mvnvFiGGCJn+JHrl11Xhvm1RffV59uJLyqRqnSK+49KRRFBc= )

SIG A 3 2 60 20001008092222 (  
  20000908092222 59140 randy.se.  
  CA+y7dFKpTEeArYQz15FhLmVJX+mZSdgvoqWcLqzwTexTQhJRTntntnc4= )
NXT

SIG NXT 3 2 900 20001008092222 ( 20000908092222 59140 randy.se.
CLbmqDnpO45UJnrYLMxoXgIIgXm0cN5Hg4DEjakw1vocE5mcAOsoLPg= )
NXT ns.randy.se. ( A NS SOA SIG KEY NXT )

and then there is NS delegation
A Classic Zone File

$TTL 60
@  SOA  randy.se.  randy.psg.com.  (  
200009080 ; serial 
15m ; refresh 
15m ; retry 
15m ; expiry 
15m ) ; minimum
NS  ns.randy.se.
A  195.149.150.111
ns A  195.149.150.111
Your Zone on DNSSEC

$ORIGIN .
$TTL 60 ; 1 minute
randy.se IN SOA randy.se. randy.psg.com. ( 200009080 ; serial 900 ; refresh (15 minutes) 900 ; retry (15 minutes) 900 ; expire (15 minutes) 900 ; minimum (15 minutes) )
SIG SOA 3 2 60 20001008092222 ( 20000908092222 59140 randy.se. CAs/Br+mvnvFiGCCJn+JHrl111Xhvm1RffV59uJLyqRqn SK+49KRRFBc= )
NS ns.randy.se.
SIG NS 3 2 60 20001008092222 ( 20000908092222 59140 randy.se. CEa9IY2yFZvKhCSfykS42vAJ8AflnFC5OpwFLIELdDxP RDSoojkaijA= )
SIG A 3 2 60 20001008092222 ( 20000908092222 59140 randy.se. CA+y7dFKpTeeArYQz15FhLmVJX+mZSdgvoqWcLqzwTex TQhJRtntnc4= )
$TTL 900 ; 15 minutes
SIG NXT 3 2 900 20001008092222 ( 20000908092222 59140 randy.se. CLbmqDnpO45UJnrYLMxoXgIIgXm0cN5Hg4DEjakwlvoc E5mcA0soLPg= )
NXT ns.randy.se. ( A NS SOA SIG KEY NXT )
$TTL 60 ; 1 minute
A 195.149.150.111
$TTL 600 ; 10 minutes
SIG KEY 1 2 600 20001007112056 (20000907112056 34309 se.
dP5pYzWKRiEZ+Is0xXhgDJZIV1IN+KEg0nzNtDQ3dnTC EZSEyibi0yB08c21/riPkqkW1i1GC+/9yacM81Gb1EL 8vbkpEprrysIEGQya3ktHc0Lu+Q+rMFf1iCspBZytGVj y1Z11nfNgGm0DRWNCoGQpPb5hCvuzjj6m0OMXIM= )

$TTL 3600 ; 1 hour
KEY 256 3 3 (CPZIU8BpUNfNyBF5Fobd7W26+rjxe9sBUTcZc0cJmim hbFjwWUM5fbUmgJEcwKleID86PP+Yg0K/QceZb3Pstap 613uOZokBjCvFhosY9LIArmmYULvo6QmizsF79GZfV bk10vDfOCZaA8ehURnTTWxz8LiikAXOouGNvDQBSX0tU epISThM4pEuPW1mjrGrL+ukU8BRk+PRb/dnVF4D8MkEvN wnLz8Bc9t+dzKrzIBodVaKCV689hQ67uA0TDHG4fZrR eWeYifAwnGuWqd9Tmt8zOrBEx1qmGebfKYCa9ecVS3c m799N8WhS9rtSlczhd3YJ15jrM0MXFdvo8NHjnNiMYte jje0K0fSiv9v6ITVty170gs+WkD4FDxJUKUVEpScF 9+R2EO5oE9SOHM62uE8sqG2PjGVEQ+BLQ/q05bEU2qxS op30y1Gn2COeh1xylJ3f4M/ikDJHooQprjFtnd5C6rhH HhFfT7wBIILwXvhW41wkoCnZjjwY9So3jm/ws6Jek7/XY Vm48aptkFD7 )

$ORIGIN randy.se.
$TTL 60 ; 1 minute
SIG A 3 3 60 200010071120222 (20000908092222 59140 randy.se.
CGMF3CS1ljXJNj1Wy8BsfWMsObS8VE8VC8uxyRZno w4PilDU/yTY=)

$TTL 900 ; 15 minutes
SIG NXT 3 3 900 200010071120222 (20000908092222 59140 randy.se.
CAcGnq6KPOGNP3Iz4bNCC2+2XY7zCcf/v/pBj7GC5Zo LWDmW2B0RqI= )

NXT randy.se. (A SIG NXT)

$TTL 60 ; 1 minute
A 195.149.150.111
Key Management

- There is no way to revoke a key, so TTLs will be short and resigning common
- All subzones have to be resigned if parent zone is resigned
- What happens when COM rolls over?
- If a child zone fails to submit new keys if old expires, name servers would not return an answer! should the parent generate a null key?
More Fun!

• NXT does not work with wildcards, and we do not know how to fix this.
• It is unclear how to sign a dynamically updated zone as changes can be being made during the signing process.
• Mechanisms are needed so registry, registrar, and registrant can establish a trusted relationship.
Workshops

- Held a few times a year to
  - Debug software
  - Develop and document operational processes
  - Find missing tools
  - Educate folk
    - DNS Admins
    - ccTLD Admins
    - NICs
Serious Problems

• Size of root response exceeds UDP MTU
• Key management (no revocation, ...), a fundamental problem with cache systems
• API complexity
• No consistent documented threat model
• Blue ribbon seminar in DC, deployable in root by EOY 2000. No DNS expertise was present.
Workshop Conclusions

• Latest workshop concluded it is possible to run a (small) secure DNS server, though not as a production service

• It was generally agreed that DNSSEC would not be deployable for at least another year. The technology is simply not mature enough, and many of the administrative issues are unresolved.

• See <http://www.centr.org/docs/technical/dnssec-ws-report.html>
The aroot Experiment

- **Many** distributed root servers
- All in the same partitioned ASn
- Uses v4 anycast
- draft-ietf-dnsop-ohta-shared-root-server-00.txt
- draft-ietf-dnsop-ohta-shared-root-server-test-00.txt
An Example (AS-paths seen at C)

- RS-A-C
- RS-D-C
- RS-F-C
- RS-G-E-C
- RS-G-E-B-C
- RS-A-C
- RS-B-C
- RS-B-E-C
- RS

: ASes (AS# = RS) containing a root server with anycast address

: Usual ASes
TSIG (RFC 2845)

- Authenticated servers/resolvers, not data
- Manually configured shared secrets
- Great for small static universes
- Might be used for dynamic update in small universes
- But does not scale
TKEY (RFC 2930)

- Key distribution for TSIG
- Pseudo-RR, not in zone files
- Allows Diffie-Hellman exchange etc.
SIG(0) (RFC 2931)

• Signs concatenation of server's response and corresponding resolver query
  • keys are of the servers, not the data
  • public keys in KEY RRs, not shared secrets
• SIG RR looks like 2535 SIG
• The goal is dynamic update security without need for full dnssec
Dynamic Updates (RFC 2136 and secure 2137)

- Wants key online! See RFC 2541
Notify (RFC 1996)

- When primary detects an update, it can tell the secondaries to ask for an AXFR
- List of secondaries is the NS RRset
- Deployed and seems to work
- Notify has significantly improved the global propagation of DNS additions
IXFR (RFC 1995)

- Incremental zone transfer
- Sends only the changes in the zone
- For distributing updates of large zones to secondaries
- Implications for resigning
- Not deployed
Controlling Extensions

- edns0 (RFC 2671)
  - OPT pseudo-RR
  - Allows UDP payloads larger than 512
  - Primitive encoding of capability levels

- edns0.5
  - draft-ietf-dnsext-edns0dot5-02.txt
  - Complex encoding of capability/extension levels
SRV (RFC 2782)

- Service Location
- Uses underscored service names

$ORIGIN foo.edu.

_ldap._tcp SRV 0 1 42 server

- priority - must use lowest reachable
- weight - prefer higher
- port - to use on server
- target - normal FQDN
NAPTR (RFCs 2168/2915)

• combines lookup and rewrite
• allows non-DNS format names
• allows DNS changes to relocate urns

```
order pref flags service regexp replacement
label NAPTR 100 50 "s"  "z3950+I2L+I2C" \
  "_z3950._tcp.gatech.edu.
```
NAPTR fields

- **order** is required order of seeking a stop-on-first-match
- **pref** allows alternative protocols etc. within ordering
- **flags** tell if the result is a
  - **s** SRV RR
  - **a** A/A6/AAAA
  - **u** URI
  - **p** protocol-specific
• +1–206–780–0431 becomes
• an E.164 number +12067800431
• and a domain name
  1.3.4.0.0.8.7.6.0.2.1.e164.arpa.
• NAPTR service E2U for E.164 to URI

$ORIGIN 1.3.4.0.0.8.7.6.0.2.1.e164.arpa.
NAPTR 1 42 "u" "sip+E2U"  
  "!^.*$!sip:phone@psg.com!" .
NAPTR 2 42 "u" "tel+E2U"  
  "!^.*$!tel:+12063107173!" .
NAPTR 99 42 "u" "mailto+E2U"
  "!^.*$!mailto:randy@psg.com!" .
NAPTAR with LDAP

- All Swedish services are in an LDAP server

\$ORIGIN 6.4.e164.arpa.

NAPTR 1 42 "u" "ldap+E2U" \ "!^+46(.*)$!ldap://ldap.se/cn=0\1!" .
Internationalization (RFC 2825)

• Protocol is 8-bit clean, but ...
• Imposing policy on syntax
• Really about Applications
• Politics, greed, and balkanization (ICANN, NSI, Semich, ...)
• IETF is on the technologic solution trail
iDNS Approaches

Per-Language Directories!

Directory Layer

ASCII DNS
(but unreadable by humans)
Maybe a Fresh Start

• Need to differentiate content from existing name space
  • so it can tell us which language
  • so we don’t break old/existing universe

• Do not need to change the base protocol

• Maybe want new RR types
A Fresh Start

$ORIGIN MY.DOMAIN.

FOOX MN A 666.42.7.11
BARRE MN MX 10 FOO
Some Consequences

• It would provide a new name space

• Protocol does not need to change

• May have new or different RR types
Another Consequence

$\text{ORIGIN ML.NET.}$

. ML NS AROOT
AROOT ML A 666.42.7.11
//anycast!

• A new root set
• Could be anycast
• So could be dnssec signed and A6 without overflowing UDP MTU
So Attend the IDNS WG and Subscribe to the list

idn@ops.ietf.org
How we Made this Camel

Applications

Security

Incremental Change

Don’t Say No

DNS

Time

Committees

IPv6

iDNS
Final Thoughts

La perfection est atteinte non quand il ne reste rien à ajouter, mais quand il ne reste rien à enlever.

You know you have achieved perfection in design, not when you have nothing more to add, but when you have nothing more to take away.

-- Antoine-Marie-Roger de Saint-Exupery
I did not omit it from the specification because I ran out of ink

-- Niklaus Wirth