DNS at MCI

Large DNS servers with many zones

Presented to the DNS WG
Presented by Andre Koopal
05/05/2004

Contents

- A little history
- Some numbers
- Our setup
- Problems we see
- How to manage customer support
- Conclusions

A little history

- NLnet, PIPEX, Iway, EUnet, INnet, UUNET
- MCI ended up with many different systems due to takeovers. We are busy consolidating, but that takes time. So we are left with different systems per country at the moment
- All had reasonable large nameservers with lots of zones for customers, both primary and secondary.
- All maintained by one team, but legacy still in place

Some numbers

Country	Prim	Sec	Servers	Queries/day
BE	12.211	1.281	2	11.803.456
DE	39.015	15.580	2(10)	41.429.393
FR	9.069	2.062	2	14.453.983
NL	24.890	7.381	2	13.733.212
UK	43.407	18.540	2(75)	105.854.887
EU	19.187	6.024	4	29.307.393
Total	147.779	50.868	14	216.582.324

Our setup

- All countries work with a system where you edit centrally
- Some systems push out changes regularly (once an hour) others use a hidden primary that get reloaded regularly together with pushing out the config regularly as well
- Reloads are done automatically during business hours, logs are tailed by a watch script to spot errors
- Possibility to restrict access to individual zones based on username or group
- Otherwise no real accesscontrol, but we do log who changes, no generic accounts
- Besides the authorative nameservers of course also a lot of resolving nameservers for customers and some specials (like TLD's).

Our setup: the software

- Authorative nameservers still tend to run bind 8
  - Bind 9 uses more memory and takes longer to start.
  - Bind 9 is more strict in checks which needs work to get the changes in the zones to get them loaded into bind 9
  - Bind 9 doesn't do autodelegation if you run both the toplevel zone and the subzone, and don't mention ns records in the toplevel zone
  - Stats of bind 9 are different, they have less data and aren't logged automatically but need to be triggered externally
  - Also bind 9 rejects secondaries based on the checks where we don't have control over the content as they are customer zones.
- For caching nameservers we mostly use bind 9 at the moment.
  - Although it uses more memory and cpu we need features that are only available in bind 9.
- Please don't use software development for politics.

Problems we see

- People lowering their serial without telling us
  - Simple solution, watch the logfiles for warnings about serial is smaller then ours, automatically remove zone and reload the one zone
- People removing zones/altering delegations we run secondary for without telling us
  - We have lots of customers running for example webfarms and running secondary for them is an add-on service
  - Problem is that the queue for zone transfers is filling up, due to which it can take much to long before a zone is transfer after scheduling
  - NX domain wouldn’t be a problem, but mostly waiting on timeouts.
  - Partly solved by by running consistency checks against the top level nameservers, but registrations can take long.
  - However, the checks also fail sometimes because tld operators block our machines because of to many queries and we don’t have access to the top level zones

Problems we see (cont.)

- Trying to rename or re-ip a nameserver.
  - Because of legacy names, or when trying to scale you sometimes need to rename nameservers.
  - We are often not the contact for secondaries, and also not always for primaries, making it almost impossible to migrate.
  - Easiest is to put a new server next to it, and migrate slowly. However resources are limited these days.
  - Any ideas for this problem from the TLD operators?
- Can somebody please remove nsp0.nl.net and ntp0.nl.net from the old Internic nameservers?

How to manage Customer Support

- In the old days at NLnet Customer support edited the zones and named.boot/named.conf directly.
- Serverops reloaded the nameserver at the end of the business day.
- Mostly twice, once to find the initial problems, second to get those active as well. But with some bad luck, you were busy fixing for an hour.
- First solution: simplify named.conf format
  - The ‘new’ bind8 config format is prone to errors if edited manually
  - Reintroduced simple format again for support to edit. Also makes introducing of standard features for zones easier.
- Second solution: provisioning script
  - We basically removed access from support on the nameservers itself and made them edit via a script on a central server doing checks.
  - Data is pushed out via rdist.

How to manage Customer Support (cont.)

- Third solution: zonecheck
  - Provisioning software runs zonecheck after editing zone.
  - Can be ignored, so people do, so we mail as well if you ignore.
  - Force FQDN's on right hand site of RR's, so you can check on forgotten trailing dots.
- Available at: http://www.slowthinkers.net/~marcel/zonecheck.html
- Fourth solution: logparsing
  - Scan logfiles on rejected zones after reload.
  - If zones rejected, mail last editor (RCS) and an alias.
  - Be strict on people ignoring these mails

Conclusions

- lots of in-house coding/scripting needed to make a reliable and redundant auth. nameserver complex
- lots of dependencies on third parties when running authorative nameservers (registrars, customers, ...)
- It would be great if there is a way to change all instances of a nameserver with a TLD operator independent of who owns the zone, at long as it is the owner of the nameserver
- Getting a copy of the TLD zones for consistency checks would be really helpful
- Not switching to bind 9 mainly due to too strict checks which can’t be switched off easily.

Questions?