DNS Services Update

Anand Buddhdev

RIPE 68 - May 2014

K-root

- Business as usual with 17 nodes
- Architectural changes coming up
    - Smaller DNS-in-a-box servers
    - Modest requirements for hosts
    - Current five global nodes will be maintained
    - Phased migration of existing local nodes to new model
    - Increased diversity with BIND, Knot and NSD

Authoritative DNS

- Primary
  - ripe.net, e164.arpa and other forward zones
  - top-level reverse zones of RIPE NCC
- Secondary
  - in-addr.arpa and ip6.arpa
  - 77 ccTLDs
  - forward and reverse zones of other RIRs
  - several miscellaneous zones (e.g. as112.net, afnog.org)
  - over 4,200 reverse zones for LIRs (/16 for v4 and /32 for v6)

Authoritative DNS

- Two active sites - Amsterdam and London
    - Peaks of 120,000 q/s
- Third site in Stockholm ready
    - Arranging transit
    - Will become active by the end of May
- Stockholm site as a backup
    - Add provisioning capability
    - Second distribution site

Secondary DNS for ccTLDs

- Reliable DNS for smaller and developing ccTLDs
- No agreements or SLAs
  - To be addressed by action item 67.1

Bound to BIND

- Until recently, only BIND was in use on the authoritative DNS cluster
  - Mature
  - Smallest memory footprint (11 GB)
  - Add/remove zones without stopping service
  - Views for separation of zones into logical servers
- Main downside
  - Entire cluster vulnerable to the same bug

Motivation for diversity

- Resilience
  - Bugs in one application cannot bring down entire cluster
- Improve software
  - Exposure to our odd mix of 5200 zones would surely tickle interesting bugs

Requirements for DNS software

- Runs on CentOS Linux
- Easy to package and deploy (RPM)
- Implements DNS and DNSSEC properly
- Runs under supervisors, such as daemontools, upstart and systemd
- Can be reconfigured without stopping service
- Zones can be added or removed without stopping service

Candidates

- Knot DNS
- NSD 4
- Nominum ANS
- BIND 10
- YADIFA

NSD 4

- Built atop NSD 3’s mature DNS code, but with new architecture
- Can add and remove zones on the fly
- Stable master process
  - Allows supervised execution
- Supports all current DNS standards
- Highly responsive team of developers

Knot DNS

- Authoritative DNS server
- Small and light-weight
- Stable master process
  - Allows supervised execution
- Supports all current DNS standards
- Highly responsive team of developers
  - Several features were added at our request

Nominum ANS

- Authoritative name server
- Supports all current DNS standards
- Will be used as a provisioning master next to BIND

BIND 10 and YADIFA

- BIND 10 was still in development
  - Not ready for production use
- YADIFA also needs more work
  - No dynamic reconfiguration
  - No NSID
  - Some bugs in notify code

Memory usage

- BIND uses about 11 GB, Knot uses about 17 GB and NSD 4 uses about 25 GB
    - NSD 4 has a “nodb” mode - uses about 17 GB
- Knot’s memory usage will go down with 1.5

Startup time

- BIND 9.9 loads all zones in about 45s
  - BIND 9.10 with the map on-disk format starts in 15s
- Knot takes about 90 seconds
- NSD 4 takes over 3 minutes
  - Zones are loaded serially

Shutdown time

- BIND 9.9 takes about 30 seconds
  - BIND 9.10 stops in about 5 seconds
- Knot takes about 40 seconds
- NSD 4 stops almost immediately

Service separation

- BIND has views
  - Allows serving some zones on their own IP addresses
- Knot and NSD 4 don’t have views
  - Run multiple instances, each with their own config file
  - Managed using upstart (no PID files, process supervision, service dependencies)

Bugs and issues

- NSD 4 returned SERVFAIL for unconfigured and expired zones
    - Changed in 4.0.3 to return REFUSED for unconfigured zones
- Several NSEC3-related bugs in both Knot DNS and NSD 4
- Knot DNS's zone parser was rather strict
- Some types of zone transfers crashed Knot and corrupted NSD's database

Managing diversity

- All software packaged into RPMs and kept in our private repository
- Ansible for configuration
  - Inventory tells server its role
  - Roles are mutually exclusive
- TSIG keys, zone list and masters stored in YAML files
- Jinja2 templates for each name server type
  - Expanded on server and filled with zone data
- Trivial to switch name server software

Questions?