KNOT DNS

Alix GUILLARD
CZ.NIC
Americká 23
12000 Praha

Ondrej Surý
CZ.NIC
Americká 23
12000 Praha

Résumé

KNOT DNS est un nouveau serveur DNS spécialement conçu pour les domaines de premier niveau et les hébergeurs. Il offre une alternative aux serveurs les plus utilisés comme BIND : les performances mesurées le place parmi les serveurs les plus rapides. Il est spécialement adapté pour les zones contenant plusieurs centaines de milliers d'enregistrements. Il est utilisé sur les serveurs faisant autorité pour le .cz (1 million de domaines) et chez des hébergeurs comme hosting90 ou igloonet. Il serait adapté aux organismes servant des grandes zones comme certaines universités ou de gros hébergeurs en France.

Mots-clefs

DNS, IPv6, DNSSEC, TLD, noms de domaine

1 Introduction

KNOT DNS est un nouveau serveur de noms (DNS) faisant autorité et spécialement conçu pour les domaines de premier niveau et les hébergeurs. Il a été développé par le registre tchèque des noms de domaine (CZ.NIC) pour répondre à ses propres besoins. Le registre CZ.NIC est l'un des 30 plus importants[1] registres nationaux (ccTLD) du monde. Sa base a dépassé le million de noms de domaine en décembre 2012 et la croissance de cette base continue à être forte[2].

CZ.NIC participe aussi aux groupes de travail de l'IETF (*Internet Engineering Task Force*) impliqués dans l'évolution des standards du DNS, aux travaux du groupe de recherche DNS-OARC ainsi qu'aux travaux du conseil des TLD européens CENTR. Ces activités permettent une implémentation rapide des nouveaux standards au sein de KNOT DNS : CZ.NIC opère, grâce à KNOT DNS, une zone de taille importante et avec des temps de réponse les plus réduits possible, tout en intégrant les standards en matière de serveurs DNS.

Cette démarche n'empêche pas de respecter l'expression des besoins des utilisateurs. Par exemple, bien que le tchèque soit une langue utilisant un alphabet avec des diacritiques, les IDN (noms de domaine internationalisés) ne sont pas acceptés sur la zone du .cz. L'acceptation de ce standard en production ayant une influence directe sur l'activité des utilisateurs, CZ.NIC sonde régulièrement ses membres sur la mise en place des IDN. Ces derniers ont toujours choisi de ne pas implémenter les IDN.

1. http://ftp.isc.org/www/survey/reports/2013/07/bynum.txt Classement TLDs hostcount
2. 8 % de progression entre novembre 2012 et octobre 2013. Source : https://stats.nic.cz/

2 Le contexte des logiciels serveurs DNS

De nombreux serveurs DNS sont disponibles sur le marché : les plus connus étant BIND, Microsoft DNS, TinyDNS et PowerDNS. Nous ne nous attarderons pas sur l'ensemble de ces logiciels[3], mais plus spécifiquement sur trois d'entre eux utilisés dans le monde des TLD : BIND, produit par l'ISC, qui est utilisé largement, NSD, disponible depuis le début des années 2000 et qui est utilisé par un des serveurs racine du DNS depuis presque 10 ans et enfin KNOT DNS disponible depuis 2011.

Dans le monde des registres Internet, la continuité de service est cruciale et les registres déploient leurs services en portant un effort particulier sur la sécurité et la redondance. Il est d'usage pour un TLD, de placer les serveurs faisant autorité sur des points du réseau fortement connectés comme les points d'échanges Internet. Ceci augmente ainsi la fiabilité du service en diversifiant les chemins d'accès (routage) entre les clients et les serveurs.

La même attention devrait être portée sur la résilience logicielle. Une zone devrait être servie par des serveurs ayant des architectures logicielles différentes. Si l'un des logiciels est victime d'une faille exploitée à grande échelle, les serveurs tournant avec un logiciel différent pourront servir la zone et assurer la continuité de service. Cette problématique a commencé à être prise en compte depuis quelques années afin notamment de limiter les failles de sécurité découvertes dans BIND, le serveur le plus utilisé. Cela a été rendu possible grâce à l'apparition d'alternatives logicielles comme NSD, puis KNOT DNS. Pour les serveurs faisant autorité, l'utilisation de plusieurs logiciels serveurs différents leur permet d'implémenter les nouveaux standards ou de mettre en place des fonctionnalités intéressantes comme dans le cadre de la mise en place de politiques lutant contre les attaques de déni de service par le DNS.

De plus ces 2 logiciels (NSD et KNOT DNS) permettent de limiter la complexité du code utilisé et donc les bugs potentiels. Seul le code implémentant les fonctionnalités de serveur faisant autorité est embarqué dans ces logiciels ce qui n'est pas le cas actuellement de la version 9.x de BIND qui est compilé sans distinction avec les fonctionnalités de serveur itératif et récursif (une séparation des deux fonctions sera possible dans la version 10.x de BIND).

3 Pourquoi KNOT DNS

3.1 Histoire du projet KNOT DNS

C'est pour offrir une plus grande diversité aux opérateurs du DNS, en particulier les TLD, que CZ.NIC a commencé à travailler sur ce projet de serveur en 2011. Le développement de KNOT DNS a alors été confié au département Recherche & Développement « CZ.NIC Labs » du registre .cz. La première présentation publique a eu lieu début novembre 2011 avec la mise à disposition de la première version Beta[4]. Dès le début, KNOT DNS s'adressait en priorité aux gestionnaires de serveurs racine et de TLD et l'accent a été mis pour développer un serveur performant pour les zones avec un grand nombre d'enregistrements.

La version finale KNOT DNS 1.0 a été publiée au printemps 2012 et les versions suivantes se sont succédées au rythme d'une mise à jour tous les six mois environ. Aujourd'hui la version disponible est la 1.3.1 ; parallèlement à l'évolution du logiciel, l'équipe de développement s'est renforcée ainsi que la base d'utilisateurs.

Les retours d'utilisation et les demandes d'amélioration se sont faits de plus en plus nombreux. La prise en compte rapide de ces retours par l'équipe de développement permet de développer la communauté.

3.2 Les points forts de KNOT DNS

Sous licence GNU GPL, KNOT DNS est ouvert et l'équipe de développement est à l'écoute des utilisateurs en mettant en ligne toute la documentation et en étudiant toute suggestion. Basée sur un code orienté objet, la structure de KNOT

3. http://mydns.bboy.net./survey/ Classement DNS Servers

4. http://ripe63.ripe.net/archives/video/197/ Lubos Slovak - Knot – DNS, a new high-performance authoritative name server

DNS est modulaire avec une architecture de type lock-free[5], sur le modèle de BIRD, l'un des logiciels de gestion de routage les plus utilisés dans le monde.

Une singularité qui le démarque des autres logiciels de sa catégorie est la possibilité de recharger une zone sans même arrêter le service ne serait-ce qu'une fraction de seconde. Cela le rend particulièrement adapté pour servir les zones volumineuses pour lesquelles le chargement peut être long. KNOT DNS s'adresse avant tout aux registres de TLD, aux gestionnaires des serveurs racine du DNS et aux hébergeurs gérant de nombreux noms de domaine.

KNOT DNS est riche et offre toutes les fonctionnalités que l'on doit attendre d'un serveur DNS qui respecte les standards. Ce logiciel est novateur par son respect des standards, y compris les dernières technologies comme DNSSEC, NSEC3, DANE ou TSIG.

A gestion performante de la mémoire lui permet de pré-calculer de nombreuses réponses et de se positionner parmi les plus serveurs DNS les plus performants. Il rivalise par exemple avec NSD souvent utilisé comme serveur secondaire. Le support de IXFR sortant, que ne propose pas NSD permet à KNOT DNS d'être utilisé comme serveur primaire tout en offrant le même niveau de performance.

4 Sous le capot

L'architecture de KNOT DNS est modulaire ce qui permet de développer rapidement de nouvelles fonctionnalités ou de mettre en place le support d'un nouveau standard. Bien qu'écrit, testé et déployé sous GNU/Linux, KNOT DNS se veut compatible avec toute plate-forme. Cette compatibilité repose essentiellement sur la bibliothèque UserspaceRCU. Pour plus d'information sur le développement de KNOT DNS, les sources du logiciel sont disponibles sur le site web[6].

4.1 Le fichier de configuration

La configuration de KNOT DNS est conçue pour être simple et facile à maintenir. Elle s'inspire du code C avec des accolades et des points virgules. Tout, depuis les interfaces, les serveurs distants, ou encore les clefs TSIG, peut être défini.

Quelques exemples de configuration :

```
system {
  # Working directory (contains zone files)
  storage "/var/lib/knot";
  # Run-time data directory
  rundir "/var/run/knot";
}

interfaces {
  # Listen on local IPv4 address
  local4 { address 192.0.2.1@53; }
  # Listen on local IPv6 address
  local6 { address 2001:DB8::1@53; }
}
```

Définition des interfaces IPv6 et IPv4 dans le fichier de configuration

5. On parle de programmation lock-free lorsque les tâches d'exécution d'un programme (*thread*) ne se bloquent pas mutuellement pour accéder à une même ressource. Voir http://preshing.com/20120612/an-introduction-to-lock-free-programming/

6. Un dépôt git propose la dernière version disponible depuis https://www.KNOT DNS.cz/pages/download.html

```
remotes {
  # Describe secondary server
  secondary { address 192.0.2.2@53; }
  # Describe primary server with TSIG key for transfers
  primary {
    address 192.0.2.3@53;
    key primary_key;
  }
}
```

Définition des serveurs distants dans le fichier de configuration

```
keys {
  # TSIG key named primary_key using algorithm hmac-sha256
  primary_key hmac-sha256 "FWdYWd4PVIFhN2g6TvK2eIccm8kxk7Sh658CIV8aGTc=";
}
```

Indication des clés TSIG dans le fichier de configuration

Certaines fonctionnalités ne sont pas standard comme la limitation du taux de réponse DNS RRL[7] (*DNS Response Rate Limiting*) invoqué au chapitre 2. Elles répondent à la demande d'utilisateurs voulant lutter contre les attaques DDoS par le DNS.

```
system {
    rate-limit 5;
    rate-limit-slip 2;
}
```

Définition du taux de questions autorisé avant rejet par le serveur (5 est considéré comme agressif) rate-limit-slip indique que les paquets seront rejetés sans réponse une fois sur deux. La deuxième fois la réponse sera TC=1 tronquée.

7. DNS Response Rate Limiting (DNS RRL) draft par Paul Vixie et Vernon Schryver http://ss.vix.su/~vixie/isc-tn-2012-1.txt

4.2 Le fichier de zone

L'édition des fichiers de zone ne diffère pas de ce qu'il est courant de voir avec d'autres serveurs DNS (Notamment BIND). Le principe d'un fichier par domaine subsiste avec l'ensemble des enregistrements et méta données qui s'y rapporte.

```
$ORIGIN example1.com.

$TTL 3600
@          SOA     dns1 hostmaster 2010111213 6h 1h 1w 1d
           NS      dns1
           NS      dns2
           MX      10 mail
dns1       A       192.0.2.1
           AAAA    2001:DB8::1
dns2       A       192.0.2.2
           AAAA    2001:DB8::2
mail       A       192.0.2.3
           AAAA    2001:DB8::3

example1.com. IN DNSKEY 256 3 5
AwEAAeQfMxs47yMjQxjUQV9UEFh6qnSeA8pNJlPadQMI3wVD1+6n+Su7
2WSHF4C2SJtVOjf88OmiOOWjm94VhXPzv/hn/815mlic1QQUi5RvIZ7+
RlwK4Gr1YAxcOKmvl6bUwHjh4XMKlDic15XVUZWtELjC4KzU3I9By5bL VXzK2mHn
```

Fichier de zone pour example.com avec quelques enregistrements simples et emplacement des clefs pour signature DNSSEC

La génération du fichier de zone signée est aujourd'hui réalisée par les outils BIND Utilities (bind-utils). Il est également prévu que KNOT DNS 1.5 fasse cette opération à la volée si les clés sont disponibles à l'endroit indiqué.

```
example1.com. 3600  IN SOA  dns1.example1.com. hostmaster.example1.com. (
                            2010111213 ; serial
                            21600      ; refresh (6 hours)
                            3600       ; retry (1 hour)
                            604800     ; expire (1 week)
                            86400      ; minimum (1 day)
                            )
              3600  RRSIG   SOA 5 2 3600 (
                            20131115121742 20131016121742 2622 example1.com.
                            yuz8kLVcVIS4weoMYlUtPgjjG4cX0AhdtMBl
                            cgdoZk+WkeuMAMQFtNdA4zAYJG7nbTnoD5/c
                            3bdFKtH4qpxzMqDjc6hn5K6Mz5YbJiV3+YTY
                            EFQRZS/jTRhoTZz3YmIDPwAe09/Q/XViaHTn
                            gs3qrqk1Gbd/6aySmrDJt+mbSXQ= )
              3600  NS      dns1.example1.com.
              3600  NS      dns2.example1.com.
              3600  RRSIG   NS 5 2 3600 (
                            20131115121742 20131016121742 2622 example1.com.
                            rDcaYa8E1Iw8LLcD9QuBmF3GsKOZVrG0+grA
                            jL5qKwBtKJZbBjjnRQqj6Qz0kogxbOehmV4f
                            QuqCiWccXVuwBHrhhLJGBRhlfeCI5Oc+/5Mv
                            IC0wSeBz2AG5q9vtK91hdN9zQy3pW8RW3P0a
                            gNVbaN9t+v2W6fc5mQepmU8mgE0= )
               3600    MX      10 mail.example1.com.
```

Extrait de fichier de zone signée, sur la base des informations dans le fichier exemple ci-dessus.

5 Performances

5.1 Temps de réponse

À la sortie de la version 1.0 en 2011, nous avons testé les logiciels de DNS les plus utilisés pour s'assurer des performances de KNOT DNS. Le test a été réalisé en laboratoire avec le logiciel dnsperf qui envoie de nombreuses requêtes DNS et en analyse les réponses. Les premiers résultats (Illustration 1) montraient des performances supérieures ou équivalentes à d'autres logiciels du marché mais surtout une meilleure tenue de la montée en charge.

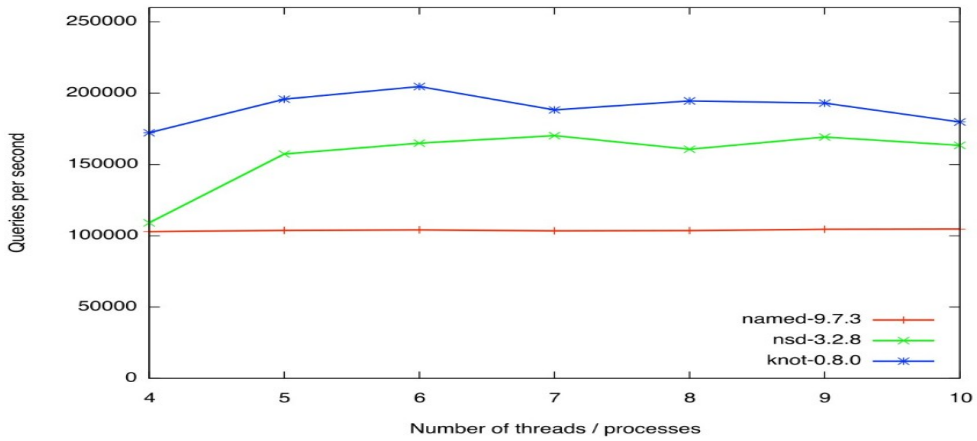

Illustration 1: Tests de performances comparant KNOT DNS, NSD, BIND (2011) avec dnsperf

Par sa nature intégrée, le logiciel dnsperf est jugée peu fiable pour mesurer dans le cas d'une montée en charge du serveur, où la machine qui opère les mesures est aussi le serveur DNS sur lequel doit être effectué ces mesures : il peut donc être fortement impacté par le logiciel dnsperf lui-même. C'est pourquoi nous avons maintenant un nouveau banc de test basé sur la solution DISTEL de NLNetlabs[8]. Cette solution permet une montée en charge du serveur DNS sans que les outils procédant à la mesure n'influent le résultat.

Les résultats sont disponibles en ligne et prennent en compte plusieurs typologies de services allant du service de type « hosting » servant un nombre important de petites zones (Illustration 2) jusqu'à un service de type « Root server » servant une zone importante et signée (Illustration 3). On y constate que la nouvelle version de KNOT DNS (1.4 alpha au moment d'écrire ces lignes) se démarque surtout des autres plates-formes dans une configuration de type « Root server ». Les tests de performance sont aussi très dépendants des configurations. Pour vous faire une idée du comportement d'un serveur, nous vous conseillons toujours de réaliser vous même ces tests de performance. La solution DISTEL étant basée sur des logiciels libres, chacun est libre de l'utiliser avec sa propre configuration.

8. Présentation de DISTEL Testlab https://www.dns-oarc.net/files/dnsops-2006/Kolkman-NSD.pdf P.28

Illustration 2: Tests de taux réponses pour une configuration DNS type « hosting » (octobre 2013) avec DISTEL source : https://www.KNOT DNS.cz/pages/benchmark.html

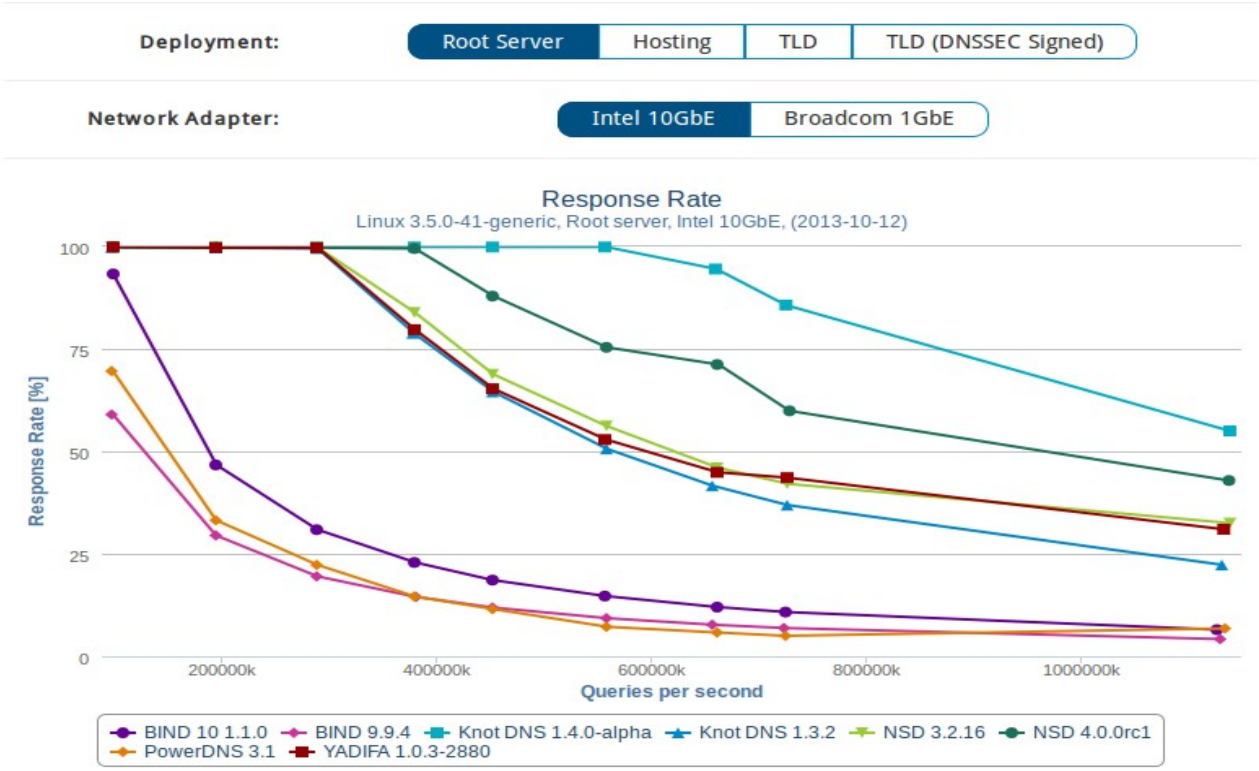

Illustration 3: Tests de taux réponses pour une configuration de type « serveur racine » (octobre 2013) avec DISTEL source : https://www.KNOT DNS.cz/pages/benchmark.html

5.2 Rapidité de chargement

La rapidité de chargement est une autre performance sur laquelle nous avons travaillé. La version 1.2 de KNOT DNS utilisait beaucoup de mémoire essentiellement en temps de compilation de chaque zone, à chaque fois que cette dernière était mise à jour. Ceci ralentissait grandement le fonctionnement de KNOT DNS[9].

Entre la version 1.2 et la version 1.3, nous avons réécrit le parseur en Ragel en optimisant les structures de données ce qui permet d'éviter la phase de précompilation des zones. En évitant cette phase, nous économisons de la mémoire et le temps de chargement de la zone reste sensiblement le même (Illustration 4). Ainsi, le serveur KNOT DNS est beaucoup plus rapide à recharger les zones qu'il sert.

Nous pensons, grâce à cette amélioration, être proche des limites théoriques de la plate-forme utilisée.

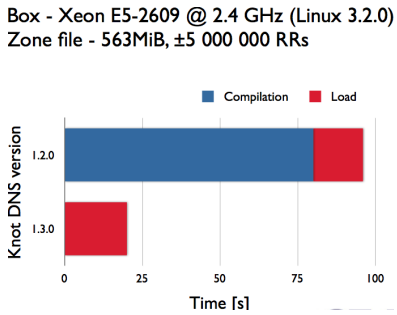

Illustration 4: Comparaison du temps de chargement entre la version 1.2 et 1.3

6 Utilisations

KNOT DNS est encore récent, il n'est pas largement utilisé mais plusieurs organisations ont fait part de leur intérêt. Certains TLD étudient actuellement son fonctionnement pour évaluer une migration possible vers cette solution. La liste ci-dessous n'est qu'indicative, KNOT DNS étant disponible librement, il est possible que d'autres organismes s'en servent sans que nous le sachions.

6.1 Services l'ayant déployé

- CZ.NIC : la zone .cz (1 million de domaines) est servie par KNOT DNS ainsi que plusieurs petites zones.
- DK Hostmaster (ccTLD du Danemark) : l'un des serveurs esclaves du .dk tourne sous KNOT DNS.
- NIX.CZ (point d'échange Internet tchèque) : l'association utilise KNOT DNS dans le cadre du projet AS112[10], qui a pour but d'intercepter des requêtes DNS envoyées sur l'Internet concernant la résolution d'adresses IP privées comme défini dans le RFC1918 : ceci permet d'améliorer les performances du système DNS public et de réduire la charge du réseau.
- Hosting90.cz et igloonet.cz (bureaux d'enregistrement tchèques) : ces deux sociétés gèrent chacune un portefeuille de 50.000 et 100.000 noms de domaine et utilisent KNOT DNS.

9. Mesures de consommation de mémoire de NSD, KNOT DNS et Yadifa par NLNetlabs https://www.nlnetlabs.nl/blog/2013/07/05/nsd-4-mem-use/

10. https://www.as112.net/ et http://www.rfc-editor.org/rfc/rfc6304.txt

6.2 Services envisageant un déploiement

- ICANN (Internet Corporation for Assigned Names and Numbers) : l'ICANN gère les serveurs racine « L » (l.root-servers.org) avec le plus gros pool de serveurs anycast (146 serveurs).
- RIPE NCC (Réseaux IP Européens Network Coordination Center) : l'organisme de gestion européen des adresses IP gère plusieurs services DNS dont le projet AS112, les serveurs racine « K » ainsi qu'un service de DNS secondaire pour les ccTLD de la région. L'équipe DNS du RIPE NCC étudie actuellement la possibilité de faire tourner ces deux derniers sous KNOT DNS.
- RU Center (ccTLD russe) : le registre russe teste KNOT DNS en laboratoire et envoie de nombreux rapports à l'équipe de développement.
- Autres ccTLD : les registres nationaux de Suède, d'Ukraine et du Royaume-Uni ont fait part de leur intérêt pour KNOT DNS.
- Autres bureaux d'enregistrement : La société tchèque Web4u examine cette solution pour servir les zones de ses clients.

6.3 Participez

Il est difficile de connaître tous ceux qui testent le logiciel sans nous en faire part. En France, il serait adapté aux organismes servant des grandes zones comme le .fr, celui des hébergeurs de grande taille et les DNS de certaines universités. Nous n'avons pas encore eu à ce jour de retours d'utilisation en France mais nous nous efforçons de renseigner ceux qui entrent en contact avec nous afin de connaître plus d'exemples de configuration, de corriger des bugs et d'apporter des amélioration utiles.

N’hésitez donc pas à installer KNOT DNS et à le tester avec votre propre environnement et vos propres données.

L'équipe de KNOT DNS est à l'écoute via les canaux suivants :

- dépôt Git : git://git.nic.cz/KNOT DNS.git ;
- KNOT DNS GitLab pour gestion des bogues : bug reporting and known issues ;
- liste de discussion des utilisateurs : KNOT DNS-users@lists.nic.cz ;
- email pour retour de bogues : KNOT DNS@labs.nic.cz.

7 Conclusion

Si vous avez lu cet article, il y a de fortes chances que vous soyez impliqués dans un réseau mettant en œuvre un service et une architecture DNS conséquente. Il y a également une grande probabilité que vous utilisiez BIND. Nous vous proposons de vous lancer et d'essayer KNOT DNS. La prise en main se fait aisément. N'hésitez pas à remonter vos remarques et de vos suggestions, ainsi que de vos problèmes ou bugs rencontrés. L'équipe de développement est à l'écoute et nous apprécions toujours les retours d'expérience sur l'utilisation de KNOT DNS. Si vous décidez de l'utiliser en production, vous aiderez à la diversification des solutions DNS : ceci permettra d'améliorer le service en le rendant plus robuste car moins vulnérable aux failles d'un seul logiciel..