Architecture DNS sécurisée, résiliente et hautement performante pour HTTP/3

1 GANDI, un acteur historique du DNS

GANDI SAS (Gestion Attribution des Noms de Domaine sur Internet) est une société française, fondée en 1999 par des précurseurs du monde Internet français avec pour objectifs de créer une alternative nationale à la gestion des noms de domaines. Aujourd'hui, l'activité de GANDI repose sur deux grands métiers :

Gestion des noms de domaines : enregistrement, renouvellement et transfert de noms de domaines. Il s'agit de l'activité historique de GANDI, un des premiers bureaux d'enregistrement (*registrar*) agréé par l'ICANN et opérationnel depuis mars 2000.

Hébergement (Hosting) : hébergement des sites internet et développement des services associés. Il s'agit d'une activité complémentaire développée dans le but d'enrichir et de compléter l'offre de GANDI.

Gandi est l'un des *registrars* français historique qui gère aujourd'hui plus de 2.5 millions de noms de domaines. À ce titre, le système de résolution des noms de domaines (DNS), brique essentielle de l'internet, est une compétence cœur de métier pour Gandi. Il permet principalement d'assurer la correspondance entre des « adresses web », connues aussi sous l'appellation de « nom de domaine », et des serveurs, référencés par des adresses IP. Actuellement, chaque requête émise par un navigateur vers une adresse web se traduit par la résolution de dizaines de noms de domaines, les sites modernes référençant de multiples ressources, retardant d'autant l'accès au contenu principal.

La capacité à tenir la montée en charge sans dégrader les temps de réponse (latences) a été un élément clef dans la conception du DNS. Ses créateurs se sont donc orientés vers une architecture distribuée hiérarchique. Composée de multiples nœuds organisés en arbre, elle répartit la charge sur les différents serveurs en utilisant la hiérarchie structurelle des données DNS.

2 Le DNS confronté à une évolution de l'utilisation d'internet

Si cette architecture patrimoniale a prouvé son efficacité pendant plusieurs décennies, elle est aujourd'hui confrontée à une évolution des usages qui accompagne une société toujours plus connectée.

2.1 De nouveaux enjeux

Gandi, comme les autres industriels du DNS, se trouve dans l'obligation de répondre à de nouvelles exigences, tant quantitatives, pour absorber les nouveaux volumes de requêtes, que qualitatives, avec des contraintes sur les temps de réponse toujours plus grandes et des garanties de sécurité et de respect de la vie privée compatible avec les exigences actuelles.

Explosion du volume de requêtes : Le nombre des requêtes DNS dépend du nombre d'utilisateurs du réseau internet et des applications qu'ils exécutent. Or, nous assistons aujourd'hui à une explosion simultanée des deux facteurs.

Si l'augmentation du nombre de personnes connectées a toujours suivi une accélération, celle-ci a longtemps été limitée par le nombre de machines. Mais la récente arrivée des objets connectés change la donne. On parle de 50 milliards d'objets en 2030 et ce sont autant de nouveaux clients pour le DNS.

Les protocoles actuels auraient peut-être pu suffire à accompagner ce mouvement au prix d'une multiplication des serveurs. Mais on assiste au même moment à une évolution des applications. Jusqu'à présent, une résolution de nom de domaine était suivie d'un traitement applicatif beaucoup plus long, ce qui limitait naturellement le nombre des requêtes. Mais aujourd'hui, ces temps de traitement séparant les requêtes sont extrêmement réduits. C'est notamment le cas avec le *cloud computing*, et notamment le FAS (*Function As a Service*), ou encore avec les micros applications embarquées dans les objets connectés. La fréquence des requêtes s'en retrouve augmenté. L'ajout de caches applicatifs locaux pourrait-être une solution, mais les TTL (*Time To Live*) sont souvent très agressif dans ce type d'utilisation.

Explosion du nombre de clients, exécutant toujours plus de requêtes, voilà l'un des défis pour les acteurs du DNS.

Contraintes temporelles Outre cette augmentation du volume, les opérateurs DNS se trouvent aussi confrontés à de nouveaux besoins en termes de délais de résolutions des noms de domaine. En effet, avec l'avènement programmé de la technologie réseau 5G, on voit émerger de nouveaux cas d'utilisation dans le domaine de la conduite autonome, des objets connectés, de la télémédecine, du divertissement, de l'automatisation de l'industrie. Le point commun entre toutes ses applications réside dans le besoin de communication à faibles latences ([18]).

L'architecture réseau *Multi-access Edge Computing* ([17, 20]) vise à répondre à cette problématique ([8]) en déployant des applications de manière distribuée sur des serveurs en bordure de réseau dans plusieurs endroits géographiques afin de fournir des ressources de calcul et de stockage pour des terminaux mobiles. Il s'agit de répondre aux requêtes des utilisateurs en utilisant les serveurs les plus proches minimisant ainsi les latences perçues.

Sécurité et vie privée : Le protocole DNS a originellement été conçu sans prendre en compte les problématiques de sécurité. Il a donc été confronté a de multiples attaques sur la disponibilité du service, la validité des réponses ou encore la vie privée des utilisateurs :

- des attaques de déni de service distribuées (DDoS) ciblant ([1]) ou utilisant le protocole DNS (attaque par amplification [21]) ont ainsi été réalisées afin de perturber voir d'interrompre le trafic internet.
- l'intégrité des données des serveurs DNS a également été victime d'attaques ayant pour objectif de rediriger les utilisateurs finaux vers des serveurs illégaux (*e.g.,* attaques par *cache poisoning*) [22]).
- la confidentialité des requêtes DNS fait également l'objet de préoccupations, l'ensemble du trafic DNS étant effectué en clair, une entité malveillante apte à capturer ledit trafic est capable de suivre et de profiler les utilisateurs ([4]).

2.2 Des solutions difficilement soutenables économiquement et écologiquement

Face à ces nouveaux défis, les industriels se tournent à la fois vers des améliorations matérielles, mais aussi vers des solutions logicielles en faisant évoluer les protocoles DNS.

Évolution matérielle : Pour répondre aux critères de latence, ils envisagent de déployer en bordure de réseau (*edge computing*) des résolveurs DNS, afin de réduire au maximum la distance séparant le client de l'information recherchée. Chacune de ces instances devra disposer de suffisamment de ressources pour faire face à un grand nombre de requêtes afin de répondre à l'augmentation prévue du trafic DNS.

Évolution logicielle : Parallèlement à cette évolution du matériel, ils devront aussi mettre en place des solutions logicielles répondant aux problématiques de sécurité et de vie privée du protocole actuel. Plusieurs propositions ont été faites ou sont à l'étude (voir état de l'art 3). Malheureusement, elles s'accompagnent toutes d'une augmentation de la charge de travail et des ressources matérielles consommées.

Qu'il s'agisse de la décentralisation ou de l'évolution du protocole, toutes les solutions envisagées pour le futur du DNS s'accompagnent d'une augmentation significative de l'utilisation des ressources matérielles. Ce qui en pratique les rend difficilement soutenables financièrement comme énergétiquement. Un travail

d'optimisation et de consolidation des ressources doit donc être effectué pour accompagner le déploiement de ces technologies. Il s'agit d'être capable de traiter le maximum de requêtes en minimisant l'utilisation des ressources, réduisant ainsi les coûts et l'empreinte énergétiques. Ce travail, détaillé dans la section 4, est l'objet de cette thèse.

3 État de l'art des nouveaux protocoles DNS

3.1 *DoT*, *DoH* : Nouveaux protocoles sécurisés

Face à ces nouvelles problématiques de sécurité, de nouveaux protocoles DNS (*DoT* [14], *DoH* [12]) réutilisant les protocoles existants de l'internet (TCP, TLS, HTTP) ont été proposés afin d'assurer la sécurisation des connexions entre les utilisateurs finaux et les résolveurs DNS.

Cependant, des travaux ([13]) indiquent que le déploiement massif de ces protocoles se traduit, d'une part, par une dégradation des performances dans des environnements réseau dégradés (3G, 4G) et d'autre part (Figure 1), par une augmentation des ressources matérielles et énergétiques incompatible avec un déploiement massif.

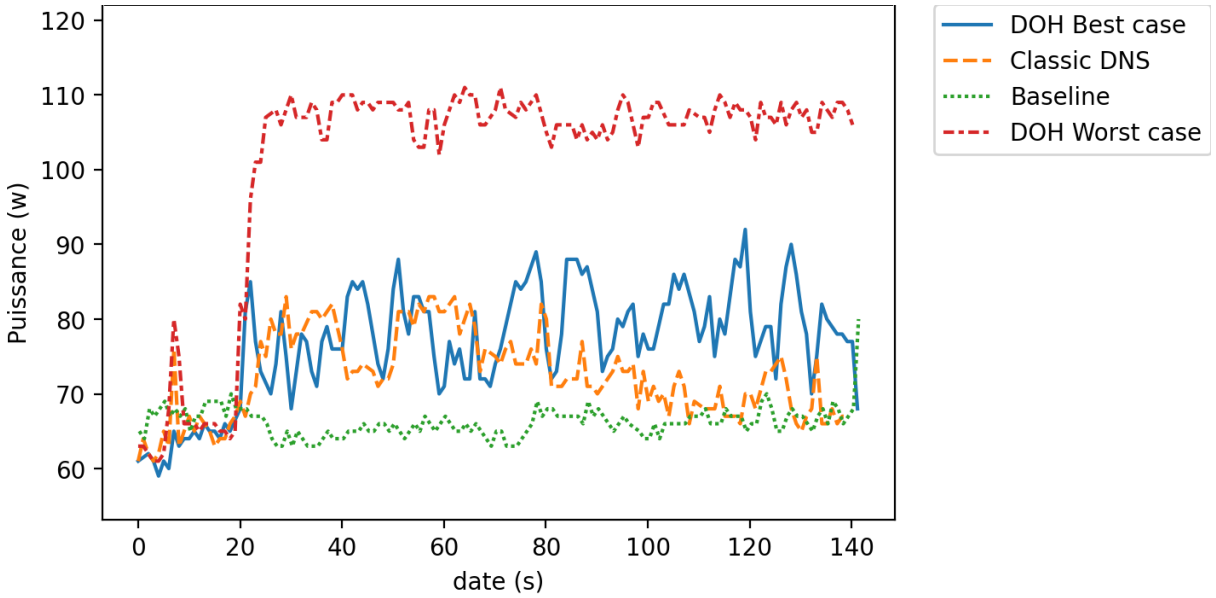

FIGURE 1 – Évolution de la consommation électrique d'un résolveur DNS selon les différents protocoles

3.2 *DoQ* : un candidat pour le futur du DNS

Un parallèle peut être réalisé entre l'évolution des besoins DNS et l'évolution de l'usage du réseau internet qui a vu l'arrivée de nouvelles problématiques (sécurité, latences) qui, initialement, n'étaient pas envisagées. Face à ces nouveaux besoins, les adaptations mises en œuvre sur les protocoles patrimoniaux (TCP, TLS, HTTP/2) se sont traduites par l'émergence de problèmes (manque de parallélisme, *head of Line blocking*, contrôle de congestion défaillant, migration de connexions à la volée impossible, latences prohibitives [9]) ne pouvant pas être résolus sans refonte globale de la stratégie réseau.

Le protocole QUIC ([16]) a donc été proposé pour remédier à ces problèmes. Il s'agit de remplacer, par un seul protocole intégré nommé QUIC, les deux protocoles que sont TCP et TLS, le premier assurant une connexion fiable entre deux ordinateurs pour acheminer sans erreur un flux de données ordonnées et le deuxième étant principalement utilisé en conjonction du premier afin d'assurer la sécurisation des données. Afin de faciliter l'adoption d'un tel protocole, les concepteurs ont choisi de se reposer sur le protocole UDP évitant ainsi la nécessité de patcher les piles réseau de l'ensemble des *middlebox* qui

assurent le routage du trafic internet. Le protocole QUIC a été sélectionné comme couche de transport par le protocole HTTP/3 ([3]).

En 2020 une nouvelle RFC ([15]) est en cours de rédaction afin de standardiser l'utilisation du protocole QUIC comme couche de transport et de sécurisation du DNS.

4 Projet de recherche

On le voit, le DNS connait et va connaitre des évolutions tant sur les protocoles sous-jacents (HTTPS, QUIC, ...) que sur son déploiement qui passe d'un modèle totalement centralisé vers des architectures partiellement décentralisées. Et si des travaux sont en cours, cela soulève encore de nombreuses questions notamment en termes de soutenabilité économique et écologique.

Dans cette thèse nous proposons donc d'aborder se problème avec la mise en place d'une plateforme expérimentale puis en travaillant, d'une part, sur le protocole et son implémentation et, d'autre part, dans la gestion des ressources nécessaire à leur exécution.

Évaluation fine du *DoQ* Le premier objectif de cette thèse vise à effectuer une étude de l'intégration du protocole DNS de type *DoQ* sur les serveurs applicatifs. En effet, ledit protocole étant encore à l'état de *draft*, il n'existe à ce jour aucun travail de recherche effectuant une étude de l'impact, que se soit au niveau applicatif (latences perçues point de vue utilisateur), qu'au niveau serveur (consommation CPU, mémoire, réseau, énergétique, capacité de passage à l'échelle...). Une telle étude est pourtant vitale pour les industriels, qui ont besoin d'évaluer le dimensionnement des futurs serveurs. De plus, en s'attachant à faire des mesures sur les différents éléments de la pile de traitement, elle ouvrira des pistes d'amélioration pour la suite des travaux de recherche. Ce travail s'appuiera sur les données fournies par Gandi afin de créer une plateforme expérimentale d'évaluation du *DoQ*.

Conception d'un *DoQ* passant à l'échelle Dans un deuxième temps, les travaux de cette thèse se concentreront sur la conception et le développement d'une solution de passage à l'échelle verticale permettant de minimiser les latences, tout en réduisant les besoins matériels nécessaires au traitement de paquets DNS. Le travail portera sur les mécanismes de *QUIC* identifiés dans la première étude de performance. Une des pistes envisagées est de mettre en œuvre des techniques de *pre-stackprocessing* en utilisant, si possible, du bytecode BPF. Le prototype réalisé sera évalué en utilisant la plateforme de test réalisé précédemment.

Utilisation des cartes programmables Pour gagner encore en latence, comme sur l'utilisation des ressources CPU du serveur, il pourrait être intéressant de réaliser un portage total ou partiel du premier prototype sur une carte réseau programmable. Cette dernière étape bénéficiera d'autant plus des travaux précédents, que nombre de ces cartes embarquent une machine virtuelle BPF ou bien, selon l'adéquation entre les besoins applicatifs et les caractéristiques des cartes réseaux, être l'occasion de développer sur de nouveaux framework ([**p4**]).

5 Une complémentarité des équipes

L'association des équipes de recherche de Gandi et du LIP6 est en réelle adéquation avec ce sujet de recherche. En effet, ce dernier se situe à la frontière du réseau et du système et demande des compétences très spécifiques : une bonne connaissance du DNS, des protocoles sous-jacents et des problématiques liées à son déploiement, mais aussi une maitrise du développement kernel, de sa pile réseau, ainsi que de la gestion des ressources matérielles (cpu, RAM, ...).

Gandi : opérateur historique du DNS, dispose de compétences à la fois sur les protocoles patrimoniaux, mais aussi de données et de métriques sur le trafic DNS existant. Fort de son service de recherche et développement, Gandi a également déjà réalisé de multiples thèses CIFRE ([10, 6, 2]) afin de réaliser de nouvelles avancées dans le domaine des réseaux, travaux qui ont été appliqués en interne pour assurer le développement de la partie hébergement de l'entreprise. Dans la dernière thèse réalisée, une solution de densification verticale a notamment été développée afin d'assurer le routage des paquets réseau sur des

cartes programmables. Ces travaux [7] se sont traduits par une multiplication par 8 du débit de routage par rapport à une solution purement basée sur Linux.

L'équipe DELYS : est une équipe mixte INRIA/CNRS. Hébergée au sein du laboratoire LIP6 de Sorbonne Université, elle a développé une forte expertise dans les développements noyau bas niveaux afin d'optimiser par le logiciel la gestion des ressources matérielles : ordonnancement dans le noyau Linux[11], gestion de mémoire dans des environnements virtualisé [19] et conteneurisé [5], ... Des travaux récents se sont attachés à améliorer l'efficacité d'applications fortement dépendantes des performances réseau en appliquant des méthodes de traitements applicatifs des paquets réseau dans le noyau avant la pile réseau tout en garantissant des propriétés de sûreté de fonctionnement. Ces méthodes se sont traduites par un gain de performances de X6 sur l'application de cache Memcached [**yoann_bmc**].

Références

[1] Abhishta Abhishta, Roland van Rijswijk-Deij et Lambert JM Nieuwenhuis. « Measuring the impact of a successful DDoS attack on the customer behaviour of managed DNS service providers ». In : *ACM SIGCOMM Computer Communication Review* 48.5 (2019), p. 70-76.

[2] Ahmed Amamou. « Isolation réseau dans un datacenter virtualisé ». Thèse de doct. Paris 6, 2013.

[3] Mike Bishop et al. « Hypertext transfer protocol version 3 (HTTP/3) ». In : *Internet Engineering Task Force, Internet-Draft draft-ietf-quic-http-27* (2020).

[4] Stephane Bortzmeyer. « DNS privacy considerations ». In : *Work in Progress, draft-ietf-dprive-problem-statement-06* 1 (2015).

[5] Damien Carver, Julien Sopena et Sebastien Monnet. « ACDC : Advanced consolidation for dynamic containers ». In : *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*. IEEE. 2017, p. 1-8.

[6] Danilo Cerović. « Architecture réseau résiliente et hautement performante pour les datacenters virtualisés ». Thèse de doct. Sorbonne université, 2019.

[7] Danilo Cerović et al. « Data plane offloading on a high-speed parallel processing architecture ». In : *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*. IEEE. 2018, p. 229-236.

[8] Batyr Charyyev, Engin Arslan et Mehmet Hadi Gunes. « Latency comparison of cloud datacenters and edge servers ». In : *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE. 2020, p. 1-6.

[9] Sarah Cook et al. « QUIC : Better for what and for whom ? » In : *2017 IEEE International Conference on Communications (ICC)*. IEEE. 2017, p. 1-6.

[10] Valentin Del Piccolo. « Isolation réseau dans un environnement Cloud Public/Hybride ». Thèse de doct. Université Pierre et Marie Curie-Paris VI, 2017.

[11] Redha Gouicem et al. « Fewer Cores, More Hertz : Leveraging High-Frequency Cores in the {OS} Scheduler for Improved Application Performance ». In : *2020 {USENIX} Annual Technical Conference ({USENIX}{ATC} 20)*. 2020, p. 435-448.

[12] Paul Hoffman et Patrick McManus. « Dns queries over https (doh) ». In : *Internet Requests for Comments, RFC Editor, RFC* 8484 (2018).

[13] Austin Hounsel et al. « Comparing the effects of dns, dot, and doh on web performance ». In : *Proceedings of The Web Conference 2020*. 2020, p. 562-572.

[14] Zi Hu et al. « Specification for dns over transport layer security (tls) ». In : *IETF RFC7858, May* (2016).

[15] Christian Huitema et al. « Specification of dns over dedicated quic connections ». In : *Internet Engineering Task Force, Internet-Draft draft-huitema-quic-dnsoquic-05* (2018).

[16] Jana Iyengar et Martin Thomson. « QUIC : A UDP-based multiplexed and secure transport ». In : *Internet Engineering Task Force, Internet-Draft draftietf-quic-transport-17* (2018).

[17] Adlen KSENTINI et Pantelis A. FRANGOUDIS. « Toward Slicing-Enabled Multi-Access Edge Computing in 5G ». In : *IEEE Network* 34.2 (2020), p. 99-105. DOI : 10.1109/MNET.001.1900261.

[18] Maria A LEMA et al. « Business case and technology analysis for 5G low latency applications ». In : *IEEE Access* 5 (2017), p. 5917-5935.

[19] Maxime LORRILLERE et al. « Puma : pooling unused memory in virtual machines for I/O intensive applications ». In : *Proceedings of the 8th ACM International Systems and Storage Conference*. 2015, p. 1-11.

[20] Quoc-Viet PHAM et al. « A survey of multi-access edge computing in 5G and beyond : Fundamentals, technology integration, and state-of-the-art ». In : *IEEE Access* 8 (2020), p. 116974-117017.

[21] José Jair SANTANNA et al. « Booters—An analysis of DDoS-as-a-service attacks ». In : *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE. 2015, p. 243-251.

[22] Sooel SON et Vitaly SHMATIKOV. « The Hitchhiker's Guide to DNS Cache Poisoning ». In : *Security and Privacy in Communication Networks*. Sous la dir. de Sushil JAJODIA et Jianying ZHOU. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, p. 466-483. ISBN : 978-3-642-16161-2.