Analyse des effets des confinements dus au Covid-19 sur l'IMRS

Bureau du directeur de la technologie de l'ICANN

Roy Arends
OCTO-008
15 avril 2020

ICANN

TABLE DES MATIERES

Ce document fait partie de la série de documents de l'OCTO. Veuillez consulter le site https://www.icann.org/resources/pages/octo-publications-2019-05-24-en pour obtenir la liste des documents compris dans la série. Si vous avez des questions ou des suggestions sur ces documents, veuillez les envoyer à octo@icann.org.

Résumé analytique

Les restrictions pendant le confinement lié au COVID-19 et les fermetures d'écoles devraient avoir un effet limité, bien que visible, sur le trafic du système des noms de domaine (DNS) hébergés dans les serveurs racine gérés par l'ICANN (IMRS). Le bureau du directeur de la technologie (OCTO) de l'ICANN a étudié l'impact du confinement imposé en France à l'échelle nationale sur les changements tant du volume que de la composition du trafic dans les quatre instances de l'IMRS en France.

Les sondes Atlas du Centre de coordination des réseaux IP européens (RIPE NCC) ont montré que le trafic des instances IMRS françaises provenait principalement de la France. Le confinement en France a débuté le 17 mars 2020 (semaine 12 de 2020). Les statistiques du trafic pour cette semaine ont montré une augmentation de 28 % par rapport à la moyenne des 6 semaines précédentes. Une analyse comparative a été effectuée entre les semaines 6 et 12 où les catégories suivantes ont été comparées :

- Requêtes pour les domaines de premier niveau existants (TLD)
- Requêtes provenant de navigateurs basés sur Chromium
- Requêtes pour les TLD de grande taille
- Requêtes pour les TLD populaires (.home, .lan, .corp et .local)
- Toutes les autres requêtes

La plupart des catégories ont montré une augmentation du trafic, ce qui a contribué à l'augmentation globale. La plus grande catégorie de requêtes provient des navigateurs Chromium, à savoir environ un tiers de toutes les requêtes reçues. Certaines catégories ont augmenté plus rapidement que d'autres. La plus forte augmentation du pourcentage est issue des quatre catégories de TLD populaires (.corp, .home, .lan et .local), qui n'existent pas. Ceci est probablement dû au plus grand nombre de personnes qui travaillent à la maison, car normalement les travailleurs sont concentrés dans des bureaux où ils utilisent un ensemble de résolveurs qui comprennent comment répondre aux domaines .corp, .lan et .local. À l'heure actuelle, les utilisateurs sont plus dispersés et lors de leur travail à la maison ils utilisent des résolveurs qui peuvent ne pas comprendre comment répondre à ces domaines. Ceci expliquerait également l'augmentation des requêtes pour .home : un plus grand nombre de personnes utilisent l'Internet plus souvent dans leurs foyers.

Les effets des confinements nationaux ont été limités, mais ils ont été visibles dans le trafic du DNS aux instances IMRS lorsqu'ils ont été observés au niveau d'un pays. Cette augmentation du trafic du DNS peut être observée au niveau global et le fait qu'aucun problème n'ait surgi suggère que l'architecture du DNS est bien adaptée pour répondre à l'augmentation du télétravail et à l'utilisation accrue dans les foyers.

1 Introduction

Les effets des confinements nationaux, des restrictions de l'activité et de la fermeture des écoles devraient être limités, mais visibles, sur le trafic du DNS aux serveurs IMRS. En général, la plupart du trafic du DNS observé au niveau de l'IMRS provient de résolveurs qui soumettent des requêtes DNS au nom de clients tels que les téléphones mobiles, les tablettes, les ordinateurs personnels (ordinateurs portables et de bureau), les consoles de jeux, etc. Ces

résolveurs ont la capacité de mettre temporairement en cache des informations, ce qui diminue la charge sur les serveurs racine. Par exemple, lorsqu'un résolveur a mis en cache des informations sur les serveurs de noms pour l'espace de noms .com, il n'est pas nécessaire de contacter les serveurs racine pour obtenir des informations sur example.com ; il n'a qu'à envoyer sa demande aux serveurs de noms .com.

Au moment de la rédaction du présent document (31 mars 2020), l'IMRS se compose de 167 instances réparties dans 83 pays. Cette étude est axée sur les quatre instances IMRS en France. La raison pour laquelle nous avons mis l'accent sur ces instances est qu'en France les fermetures d'écoles, les restrictions des activités et le confinement à l'échelle nationale ont été annoncés successivement par le gouvernement. Le 12 mars, le gouvernement a annoncé que les écoles et les universités seraient fermées à partir du lundi 16 mars. Le 13 mars, les réunions de plus de 100 personnes ont été interdites. Le 14 mars, la fermeture de tous les lieux publics non essentiels, y compris les restaurants, les cafés, les cinémas et les discothèques, a été ordonnée. Le 16 mars, un confinement à l'échelle nationale a été ordonné à partir du lendemain.

Le trafic vers les serveurs IMRS provient d'une grande variété de sources qui ne résident pas nécessairement dans le même pays que les instances faisant l'objet d'une requête. L'utilisation des sondes RIPE Atlas[1] comme proxy pour les clients du résolveur nous permet de visualiser les sondes individuelles utilisées par les quatre instances IMRS actuellement situées en France.

[1] RIPE Atlas est une plateforme de mesure de l'Internet global, ouvert et distribuéhttps://en.wikipedia.org/wiki/Internet, composée de centaines d'appareils de mesure qui mesurent la connectivité Internet en temps réel.

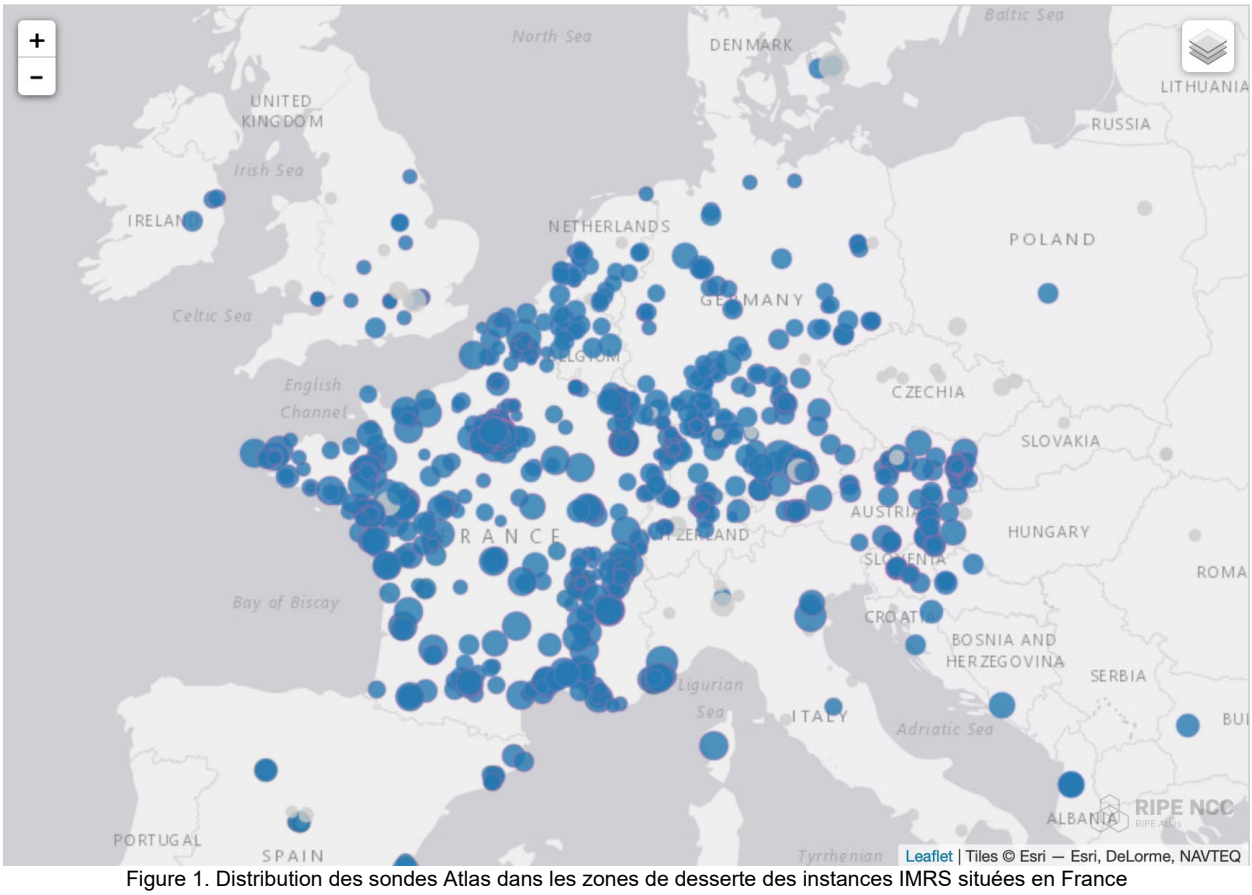

Figure 1. Distribution des sondes Atlas dans les zones de desserte des instances IMRS situées en France

Comme le montre la figure 1, malgré la bonne quantité de sondes situées en dehors de la France qui ont vu une réponse des instances IMRS susmentionnées, une proportion importante du trafic pour les instances IMRS en France provient de la France.

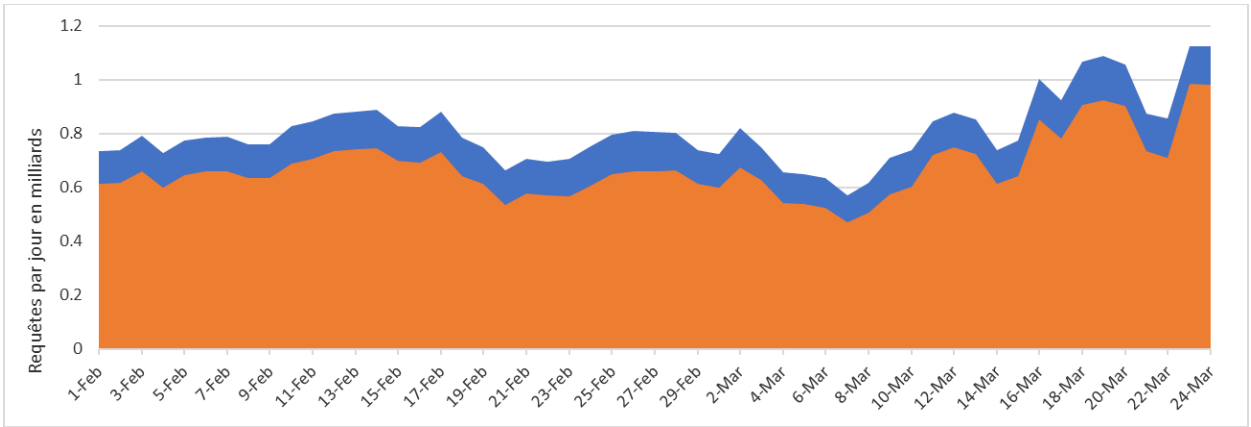

Figure 2. Volume de requêtes quotidiennes (bleu) et volume de réponses quotidiennes NXDOMAIN (orange) observés dans les 4 instances IMRS en France. La ligne noire verticale indique le début du 17 mars.

Comme le montre la figure 2, le volume du trafic a augmenté après le 16 mars. Pour comprendre ce qui a entraîné cette augmentation, nous avons examiné la composition du trafic. Nous comparerons la composition avant et après le 16 mars pour voir si nous pouvons établir une corrélation entre les changements de composition et les confinements.

2 Méthodologie

Nous comparerons deux semaines de trafic. La première semaine de février (semaine 6, à partir du 3 février) contre la semaine du 16 mars (semaine 12) qui était la première semaine du confinement. Nous classifierons ensuite les différentes tranches de ce trafic et montrerons quelle est la classification ayant connu les changements les plus importants.

2.1 Classification

Le trafic est groupé en plusieurs catégories en fonction du TLD demandé :

- **Existants :** Requêtes pour les TLD actuellement délégués à partir de la zone racine
- **Chrome** : Requêtes pour les TLD inexistants d'entre 7 et 15 caractères de long.
- **Jumbo** : Requêtes pour les TLD inexistants de plus de 15 caractères
- **.home** : Requêtes pour les domaines qui se terminent par .home
- **.lan** : Requêtes pour les domaines qui se terminent par .lan
- **.local** : Requêtes pour les domaines qui se terminent par .local
- **.Corp** : Requêtes pour les domaines qui se terminent par .corp
- **Autres** : Requêtes pour tous les autres domaines

2.1.1 Requêtes Chrome

Le navigateur Web Chromium et ses dérivés (tels que Google Chrome, les versions récentes de Microsoft Edge, Amazon Silk, et le navigateur Web d'Opera) émettent trois requêtes DNS avec une étiquette aléatoire pour détecter si le résolveur utilisé sur le réseau local redirige les domaines inexistants, par exemple, si la requête renvoie l'adresse d'un site Web de recherche « d'aide » pour les domaines qui n'existent pas. L'étiquette se compose de lettres aléatoires et comporte entre 7 et 15 caractères.[2] Étant donné que le domaine interrogé est aléatoire, le résolveur de réception ne l'aura pas mis en cache et émettra une requête à un serveur racine. Dans les réseaux sans redirection, la réponse attendue de cette requête aléatoire serait un code d'erreur NXDOMAIN.

[2] « Nous générons un nom d'hôte aléatoire d'entre 7 et 15 caractères ». https://chromium.googlesource.com/chromium/src/+/master/chrome/browser/intranet_redirect_detector.cc#150

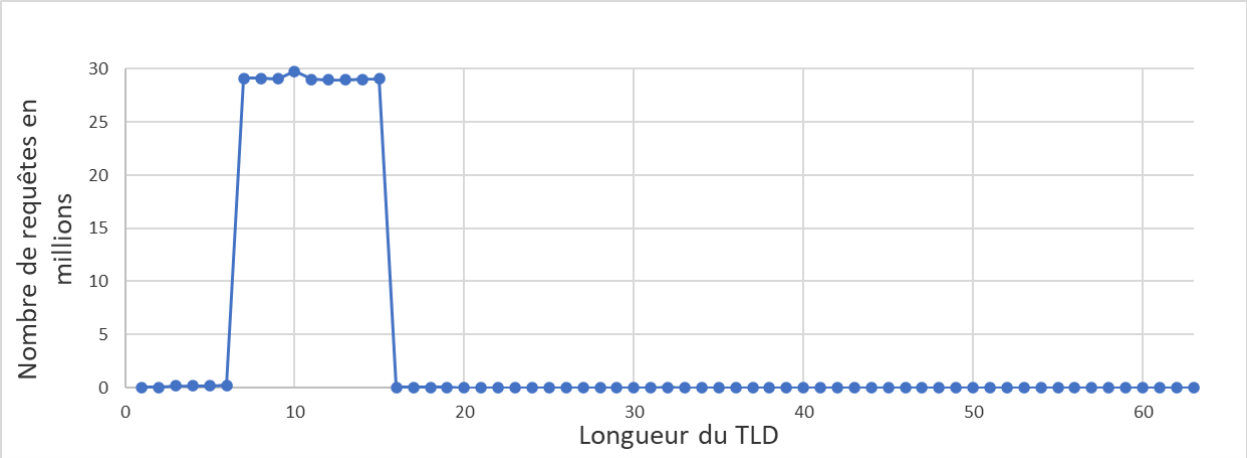

Figure 3. Histogramme du nombre de requêtes pour les TLD inexistants par longueur de TLD.

L'histogramme de la figure 3, qui affiche les données du 19 mars, montre la distribution de la fréquence des requêtes par longueur de TLD. La plupart de ces requêtes concernent des noms de domaine compris entre 7 et 15 caractères. La figure 5 montre que ces requêtes Chrome représentent 28 % de toutes les requêtes pour des domaines inexistants.

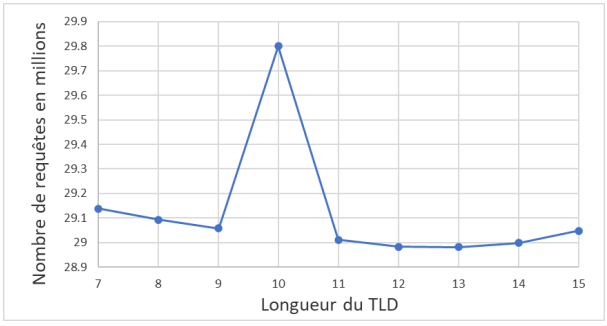

Figure 4. Détails de l'histogramme du nombre de requêtes pour les TLD inexistants par longueur de TLD.

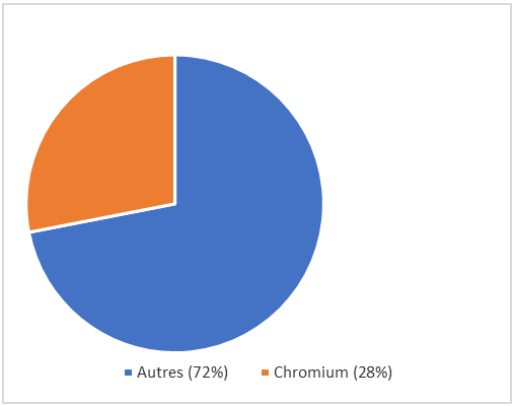

Figure 5. Nombre de requêtes Chromium sur toutes les requêtes pour des domaine inexistants.

A part les TLD à 10 caractères, la distribution entre les TLD à 7 et 15 caractères est assez uniforme. L'anomalie pour les étiquettes de 10 caractères peut être attribuée à des versions plus anciennes de Chrome qui ont renvoyé des domaines aléatoires de 10 caractères.[3]

2.1.2 Requêtes Jumbo

Ce sont les requêtes pour les TLD inexistants de plus de 15 caractères. Nous ignorons les sources ou les causes de ces requêtes.

2.1.3 TLD populaires inexistants

[3] « Faire varier la longueur des noms pour la détection du détournement du DNS ». https://src.chromium.org/viewvc/chrome?view=revision&revision=249013

Il existe une gamme d'étiquettes populaires qui n'ont pas été déléguées à la racine et qui n'existent pas dans l'espace de noms publics du DNS. Parmi les plus populaires de ces TLD inexistants on peut citer .home, .lan, .corp et .local. Ces TLD sont classés individuellement car ils ont tous augmenté en volume pendant notre étude.

2.1.4 Autres

Cette catégorie capture toutes les requêtes qui ne peuvent pas être classées dans l'une des autres catégories décrites précédemment.

3 Observations

Les quatre instances IMRS en France ont reçu en moyenne 5,4 milliards de requêtes DNS par semaine entre les semaines 6 et 11 (voir la figure 6). Les mêmes instances ont reçu 6,9 milliards de requêtes DNS au cours de la semaine 12. Cela représente une augmentation de 28 % du trafic vers ces quatre nœuds IMRS.

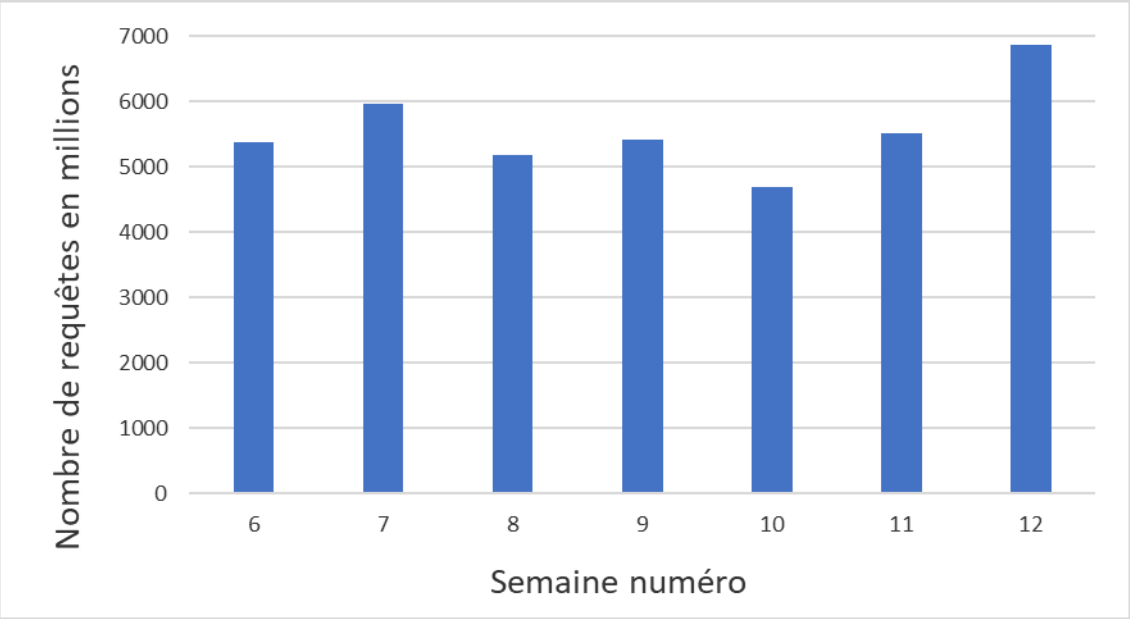

Figure 6 : Charge de requête par semaine sur les 4 instances en France, de la semaine 6 à la semaine 12.

Nous avons identifié certaines anomalies, telles que les brèves rafales de trafic ou les pannes d'entretien d'une instance au cours de cette période, mais celles-ci ont eu tendance à être de courte durée et nous ne pensons pas que leur influence soit significative sur le trafic global. D'autres modèles de trafic, tels que les modèles diurnes ou les week-ends, sont également absorbés étant donné que le trafic a été accumulé sur une semaine. Nous ignorons tout autre changement ou événement sur cette période qui aurait une influence significative sur le volume des requêtes DNS.

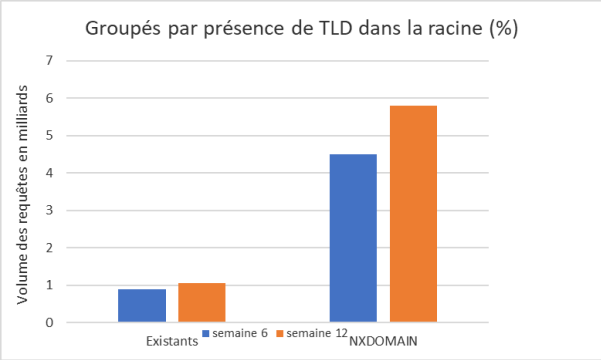

Figure 7 : Volume du trafic pour les TLD existants et inexistants dans les semaines 6 et 12

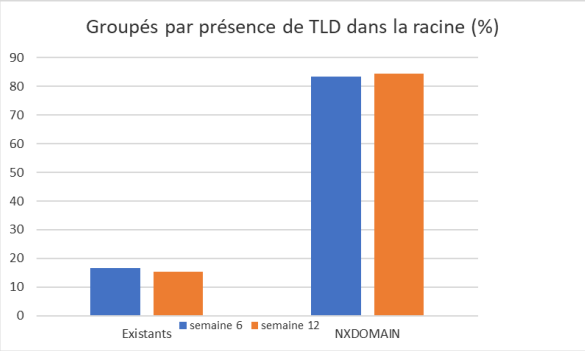

Figure 8 : Volume du trafic pour les TLD existants et inexistants dans les semaines 6 et 12 comme pourcentage du volume total pour ces semaines

La figure 7 illustre la différence de volume de requêtes pour les domaines existants et inexistants en chiffres absolus. Les deux groupes ont augmenté en volume. La figure 8 montre également qu'il y a un petit changement dans la composition du trafic, puisque le pourcentage de requêtes pour les TLD existants a diminué par rapport à celui des domaines inexistants. L'augmentation du trafic se trouve principalement dans les requêtes pour les domaines inexistants.

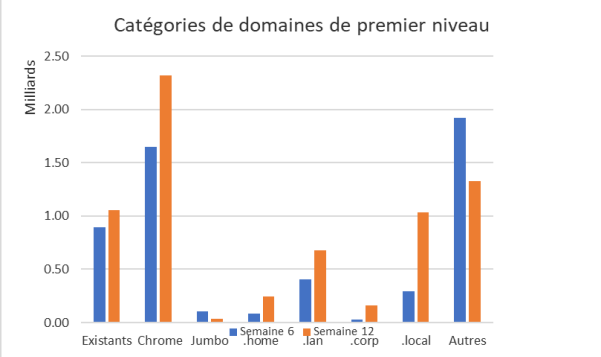

Figure 9 : Répartition du trafic dans différentes catégories, en comparant les semaines 6 et 12 en chiffres absolus.

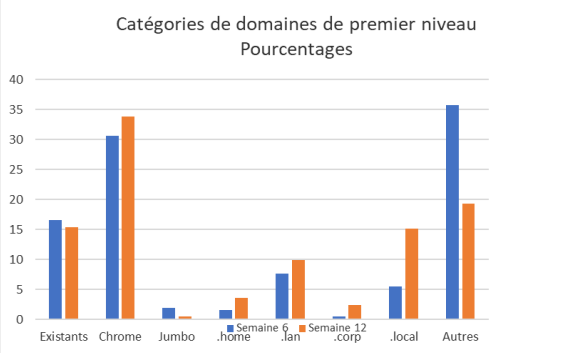

Figure 10 : Répartition du trafic dans différentes catégories, en comparant les semaines 6 et 12 en pourcentage du volume total.

3.1 Requêtes Chromium

Nous avons observé que 31 % des requêtes reçues au cours de la semaine 6 et 34 % au cours de la semaine 12 relèvent de la catégorie des demandes DNS à Chromium. Cela reste une partie importante du trafic global. Après le confinement, le nombre total de demandes a augmenté de 28 %, tandis que la partie Chromium de ces requêtes a augmenté de 41 %. Cette augmentation est probablement due à un plus grand nombre de dispositifs en ligne en raison du mandat de rester à la maison.

En raison des requêtes de détection de redirection DNS de Chromium, il y aura un taux plus élevé de requêtes DNS pour des chaînes aléatoires de 7 à 15 caractères visibles dans le trafic lorsque plus de dispositifs ayant des navigateurs basés sur Chromium seront mis en ligne. Notez que les requêtes pour les demandes DNS Chromium n'ont pas augmenté du même

pourcentage que le trafic global. Cela indique un léger changement dans la composition du trafic global. D'autres catégories ont vu une augmentation plus élevée que celle des requêtes Chromium.

Les requêtes Chromium sont la principale motivation des requêtes aux serveurs racine. D'autres instances IMRS reçoivent souvent plus de 50 % de toutes les requêtes entrantes de Chromium. Le but de ces requêtes est de vérifier si Chromium est derrière un portail captif. L'approvisionnement des serveurs racine est souvent en fonction de la charge globale sur les serveurs racine pour répondre aux besoins de mise à l'échelle. Bien que ces requêtes soient gratuites pour Chromium, le coût d'approvisionnement des instances du serveur racine ne l'est pas. Google a été informé de ce problème qui reste toujours en suspens.[4]

3.2 Requêtes Jumbo

Nous avons observé que le volume des requêtes pour les domaines TLD longs (plus de 15 caractères) a diminué. La raison de cette baisse du trafic n'a pas été examinée.

3.3 TLD populaires inexistants

Les quatre TLD les plus populaires qui ont connu une augmentation du volume ont été .corp, .home, .lan et .local. Parmi eux, .corp, .lan et .local ont connu la plus forte augmentation. Cela est probablement dû au plus grand nombre de personnes qui travaillent à la maison. Normalement les travailleurs sont concentrés dans des bureaux où ils utilisent un ensemble de résolveurs qui comprennent comment répondre aux domaines .corp, .lan et .local. À l'heure actuelle les utilisateurs sont plus dispersés et pour leur travail à la maison utilisent des résolveurs qui peuvent ne pas comprendre comment répondre à ces domaines. Ceci expliquerait également l'augmentation des requêtes pour .home : un plus grand nombre de personnes utilisent l'Internet plus souvent dans leurs foyers.

4 Conclusion

Les effets des confinements nationaux pour maîtriser la pandémie mondiale ont été limités mais ils ont été visibles dans le trafic du DNS aux instances IMRS lorsqu'ils ont été observés au niveau d'un pays. Cette augmentation du trafic DNS peut être observée en général. Le fait qu'aucun problème n'ait été constaté suggère que l'architecture du DNS est bien adaptée aux scénarios de travail à distance et à l'utilisation accrue dans les foyers

Auteurs : Adiel Akplogan, Roy Arends, David Conrad, Alain Durand, Paul Hoffman, David Huberman, Matt Larson, Sion Lloyd, Terry Manderson, David Soltero, Samaneh Tajalizadehkhoob, Mauricio Vergara Ereche.

[4] Les trois sondes aléatoires du détecteur de redirection Intranet n'ont pas de TLD et atteignent ainsi les serveurs racine.
https://bugs.chromium.org/p/chromium/issues/detail?id=946450&q=intranet%20redirect&can=2