P1: Analyse d'applications réseaux - ShadowDrive 8

Arthur De Neyer Axel Sneessens

***Abstract*—Ce rapport offre une vue détaillée sur les protocoles utilisés et les différents mécanismes réseau mis en place sur `drive.shadow.tech`. Ce drive, créé par l'entreprise française de Cloud Computing Shadow, offre un stockage sécurisé et gratuit. [1]**

INTRODUCTION

Ce rapport contient l'analyse réseau[1] d'une application: **ShadowDrive**. À travers les sections de ce document, nous énumérerons les fonctionnalités de l'application, des scénarios d'analyse spécifiques sur lesquelles nous porterons nos observations, ainsi que de nombreuses réflexions et interprétations sur ces observations. Nous nous sommes basés sur d'autres travaux du même registre afin d'assurer un résultat optimal [2].

I. FONCTIONNALITÉS ET SCÉNARIOS

A. Fonctionnalités

Notre application web est **ShadowDrive**. Celle-ci a une version web, une version mobile et une version bureau. Dans cette analyse, nous testerons la version web uniquement, étant donné que la version bureau n'est qu'un synchronisateur vers la version web, et ne contient qu'un petit éventail des fonctionnalités.
Notre application comporte les fonctionnalités classiques d'un drive:

- Créer un compte Shadow, permettant d'avoir 20GO gratuit de stockage en ligne
- Depuis son Drive, voir ses fichiers/dossiers
  - Créer · Ajouter · Partager · Ajouter aux favoris · Voir les détails · Renommer · Déplacer/Copier · Éditer · Télécharger · Supprimer des dossiers/fichiers.

Il y a aussi une partie axée sur les photos et librairies, mais nous nous concentrerons principalement sur les dossiers/fichiers.

B. Scénarios

Voici les scénarios sur lesquels se baseront nos analyses :

1) La création d'un Drive en utilisant un compte Google.
2) La connexion au Drive, suivi d'un accès à un fichier
3) L'accès le plus simple possible au drive, en étant déjà connecté, sans rien faire d'autre
4) L'ajout d'un fichier au drive, en y étant déjà connecté.
5) L'ajout d'un gros fichier au drive, en y étant déjà connecté
6) Le téléchargement d'un gros fichier, en y étant déjà connecté

[1]Github: https://github.com/GaecKo/LINFO1341-P1

C. Outils utilisés

Ce travail s'est basé sur une version propre du navigateur Chrome et Brave, avec le moteur de recherche Google et DuckDuckGo. C'est aussi sur ce navigateur que nous avons pu analyser certains éléments du réseau grâce à l'outil *Inspect* → *Network*. Nous avons aussi utilisé `Wireshark` pour la capture de paquets et de nombreuses commandes terminales tel que `dig`, `whois`, et `traceroute` pour avoir de plus amples informations.

II. DNS

A. Domaine principal

Afin de trouver les adresses IP utilisés par `drive.shadow.tech`, utilisons `dig A drive.shadow.tech`. Cette requête nous cite ces 2 adresses IPv4 (A) non-autoritaires :

- 46.105.132.157
- 46.105.132.156

Le TTL (*Time-To-Live*) pour `shadow.drive.tech` DNS A record est de 214 secondes, ce qui est relativement court. Cela pourrait être utile dans le cas où les enregistrements DNS changent régulièrement, mais augmente donc la charge de travail des résolveurs qui doivent interroger plus fréquemment les serveurs DNS autoritaires pour obtenir les mises à jour.
Notre site n'utilise étonnamment aucune adresse IPv6 (AAAA), ce qui pourrait entraîner des problèmes de connexion à l'avenir du à la pénurie d'adresses IPv4, ou forcer l'utilisation (lente) d'un Tunneling IPv6 → IPv4 [3].
Voici les DNS autoritaires:

- candy.ns.cloudflare.com
- dns.cloudflare.com

Ces deux DNS autoritaires ont 797 et 300 secondes de TTL respectivement. Cela semble relativement court, surtout en prenant en compte le fait que ces DNS ont l'air d'être localisé aux USA. Cela signifie donc de devoir actualisé régulièrement une demande vers les USA. Cela ne semble pas parfaitement optimal.
Comme on peut le voir, le fournisseur DNS de ShadowDrive est Cloudflare. Analysons maintenant ces résultats avec `traceroute` et `whois`. Lorsque nous utilisons `traceroute -U -p 443 --sport=4000 drive.shadow.tech`, nous voyons un départ depuis notre modem vers **BELGACOM**, géré par **Proximus**, à Bruxelles. Nous passons ensuite par un autre service **BELGACOM/Proximus**, toujours à Bruxelles. Nous avons ensuite 5 passages "cachés" [4], et finissons par **OVH**, en France, sans doute le CDN de ShadowDrive. On apprend

d'ailleurs que le service ShadowDrive se base sur **hubiC**, service racheté par OVH et devenu Shadow SA en 2021 [1]. Nous voyons le trajet de notre paquet à la figure 1. Pour transmettre un paquet de 84 bytes, headers compris, la commande ping nous indique un temps moyen (sur 10 essais) de 62 ms.

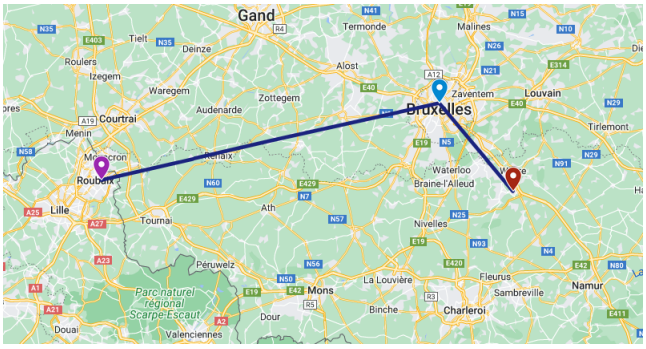

Fig. 1. Parcours d'un paquet

B. Création d'un Drive par compte *Google*

Lorsque nous souhaitons créer un Drive par un compte Google, nous utilisons de nombreux services Google. En 26 secondes de capture, nous avons 122 requêtes DNS dont 110 contenants le mot-clé Google, dont 58 `HTTPS`, 8 `AAAA` et 44 `A`. Dans ces requêtes, nous en avons certaines créées par le navigateur Google Chrome (Notamment par exemple vers `history.google.com` CNAME de `history.l.google.com`, qui sert à synchroniser l'historique avec le compte Google connecté). Nous avons cependant un grand nombre d'autres requêtes dont il est compliqué de comprendre l'origine. En voici une petite liste non-exhaustive:

- `lh3.googleusercontent.com` CNAME de `googlehosted.l.googleusercontent.com`, utilisé pour les services cloud basé sur Google, mais pas par Google [5].
- `mtalk.google.com` CNAME de `mobile-gtalk.l.google.com`, utilisé pour Firebase Cloud Messaging [6].
- `clients4.google.com` CNAME de `clients.l.google.com`, malheureusement nous n'avons pas trouvé d'explications sur ce que représentait ce service.
- `play.google.com`, correspondant au play store de Google.
- `content-autofill.googleapis.com` pour l'auto complétion de forms, ...
- `www.gstatic.com`, dont le nom de serveur autoritaire est `ns1.google.com`, correspond à un service permettant la rapidité de Google, une sorte de service de cache [7].

Ces requêtes DNS semblent parfois assez particulières et indiquent un grand nombre d'opérations effectuées par Google en arrière-plan. Il est parfois difficile de savoir ce qu'il se passe à 100% derrière ces opérations, ce qui n'est pas rassurant. Étant donné le scénario de la capture utilisant un compte Google, nous avons aussi vu d'autres requête DNS vers Google, mais cette fois-ci il est plus simple de comprendre leurs origines et utilisations. Par exemple, voici une suite de paquets dans l'ordre (simplifiée) :

1) `drive.shadow.tech` → accès vers le drive
2) `auth.eu.shadow.tech` → vérification de sécurité Cloudflare
3) `www.googleapis.com` → API vers les comptes Google
4) `accounts.youtube.com` → lien avec le compte Google (ici en passant par Youtube, ce qui est étrange)
5) `auth.eu.shadow.tech` → retour vers Shadow-Drive
6) `cdn.builder.io` → semble être l'accès vers la fin de la création du compte, vers le contenu de la page en étant connecté. Cela semble particulier mais des requêtes vers CloudFront sont éffectuées (CDN de builder.io), ce qui confirme l'utilisation de ce système.

C'est uniquement vers la fin de la trace que nous avons une requête vers `accounts.google.com`, qui est sans doute pour synchroniser à nouveau les données avec le compte Google.

Outre ces nombreuses requêtes DNS vers Google, nous n'avons remarqué aucune autre particularité dans les requêtes en tant que tel.

C. Connexion au Drive par compte *UCLouvain*

Dans ce scénario, nous avons simplement accédé au Drive avec un compte utilisant une adresse email UCLouvain. Sur 122 requêtes DNS, nous avons observé 36 IPv4 (A), 40 IPv6 (AAAA) et 46 HTTPS. Parmi ces 122 requêtes, 109 sont, étonnamment, des requêtes Google. Les autres sont des requêtes directes au Drive. Certaines de ces requêtes sont, à nouveau, parfois faites avant l'accès au site même, donc générée par le navigateur Chrome. Penchons-nous sur les différentes requêtes DNS qui semblent importantes, que nous n'avons pas encore vu :

- `google-ohttp-relay-safebrowsing.fastly-edge.com` → domaine utilisé dans le cadre de Privacy Sandbox, une initiative qui a pour but de permettre la publicité ciblée. [8].

- `passwordsleakcheck-pa.googleapis.com` → semble être utilisé par Google pour fournir un service de vérification de potentielle fuite de mots de passe. [9]

Par la suite, les requêtes DNS semblent indiquer une procédure d'accès classique à un Drive, en ne passant cette fois-ci pas par un compte Google. Nous gardons cependant un grand nombre d'étonnantes requêtes vers des services Google. Nous avons donc décidé de retenter le scénario en passant par un autre navigateur: **Brave**, et comme moteur de recherche **DuckDuckGo**. En effectuant le même scénario,

nous obtenons **0** requête DNS vers des services Google. Nous obtenons cependant 52 requêtes contenant **brave** dans le `dns.qry.name`, utilisé par exemple pour les mises à jour du navigateur. Nous observons donc bien l'omniprésence de Google lors de l'utilisation de Chrome (qui est d'ailleurs le navigateur le plus utilisé en 2023 [13]), ce qui pousse à la réflexion quant à la manière dont nos données pourraient être traitées.

D. Ajout d'un fichier dans le Drive

Dans ce scénario, en ne prenant pas en compte les requêtes vers des services Google, nous obtenons 24 requêtes DNS, dont 8 HTTPS, 8 AAAA et 8 A, toute vers soi `drive.shadow.tech` ou `auth.eu.shadow.tech`. Nous n'avons donc pas de changement dans nos requêtes DNS lors d'un upload. Nous en apprendrons plus sur le transfert de fichier en tant que tel dans la partie IV-C.

III. Couche réseau

Après de profondes analyses de nos traces, nous n'avons pas trouvé d'utilisation claire de NAT lors de l'utilisation d'IPv4. De même, cela ne paraît pas anormal puisqu'il n'y a à priori pas de communications directes entre deux ordinateurs, étant donné que tout est relié et synchronisé par l'intermédiaire d'un serveur.
En analysant les conversations IPv4 de notre analyse de paquet lors d'envois de données, nous remarquons assez clairement que l'adresse IPv4 utilisée pour réceptionner nos données est `46.105.132.156`. Cette adresse est bien la même que celle obtenue au départ de notre rapport. C'est la seule adresse que nous avons trouvée lors d'upload de données. Nous en apprendrons plus dans la partie IV-C.

IV. Couche Transport

A. HTTP

ShadowDrive se base sur le protocol HTTPs version 2.0, avec TLS 1.2 et 1.3 (plus d'informations sur TLS: VI)

B. Chargement Initial HTTP

Lors du chargement initial de la page, nous observons 140 requêtes (voir figure 2), dont environ 2/3 sont utilisées pour du CSS et du JS. Après analyse, nous avons un grand nombre de fichiers JS s'occupant de toute la partie logique de l'application (synchronisation, upload des modifications, ...). Par exemple un fichier JS nommé `merged-index.js` d'environ 14 milles lignes, appelées lors d'un upload de fichier, permet à priori de diviser un fichier à upload en plusieurs sous-parties, pour ensuite le fusionner à l'arrivée. Nous en apprendrons plus dans la sous partie IV-C. Cela explique donc ce grand nombre de chargements de fichiers JS.

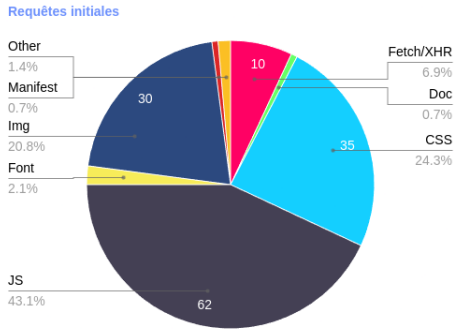

Fig. 2. Requêtes Initiales

C. TCP

Pour tester le fonctionnement des transports TCP, voici deux scénarios intéressants:

1) L'ajout d'un gros fichier au drive (153.6 MB),
2) Le téléchargement de ce même gros fichier depuis le drive.

Les statistiques sur le protocole TCP nous indiquent 4 conversations TCP vers `46.105.132.156` (voir figure 3) lors de l'upload d'un gros fichier. L'une d'entre elle se démarque par son haut nombre de bytes transférés, et est donc celle par lequel le transfert s'est fait. On remarque d'ailleurs que la majorité des paquets ont une taille de 1408 Bytes, ce qui indiquerait le taux auquel le transfert s'effectue. Les 4 conversations se finissent de manière propre. Voici un graphe intéressant sur le nombre de paquets par secondes pendant le transfert, indiquant la stabilité du réseau, ainsi que les erreurs observées (voir figure 4)

192.168.0.144	46518	46.105.132.156	443	52	34 kB
192.168.0.144	46522	46.105.132.156	443	58	23 kB
192.168.0.144	46664	46.105.132.156	443	164,377	166 MB
192.168.0.144	55632	46.105.132.156	443	23	6 kB

Fig. 3. Conversations TCP vers `drive.shadow.tech`

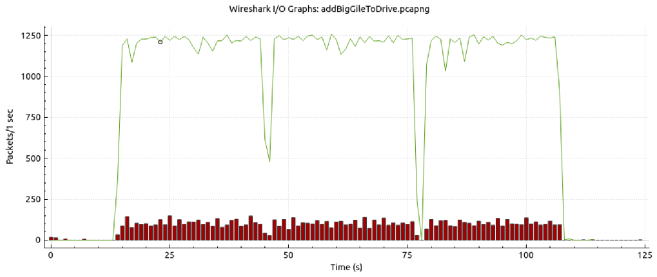

Fig. 4. Paquets/s: protocole TCP lors d'un upload

Lors du téléchargement d'un gros fichier, nous observons cette fois-ci 3 conversations TCP vers `46.105.132.157` (voir figure 5), qui n'est pas la même que celle vue précédemment, mais appartient bien aussi à `drive.shadow.tech`. Nous avons aussi un nouveau nom de domaine utilisé: `download.drive.shadow.tech`,

sans doute à l'origine du téléchargement. À nouveau, il est simple de voir la conversation correspondant au transfert du fichier en tant que tel. Les 3 conversations se terminent de manière normales aussi.

192.168.0.144	43692	46.105.132.157	443	22	9 kB
192.168.0.144	43696	46.105.132.157	443	37	16 kB
192.168.0.144	43698	46.105.132.157	443	49,106	164 MB

Fig. 5. Conversations TCP depuis `drive.shadow.tech`

Toutes les requêtes, dans les deux scénarios, s'effectuent sur le port 443, correspondant bien au port standard réservé aux connexions web sécurisées.

D. QUIC

QUIC (Quick UDP Internet Connections) est un protocole créé par Google qui permet l'envoi de paquets via le protocole UDP avec un temps de transport réduit ainsi que des connexions de multiplexage. [12]

Nous avons observé 3 paquets différents:

- "Initial" : Ce sont les paquets utilisés pour établir une nouvelle connexion en transmettant les informations pour configurer la connexion.
- "Handshake" : Ces paquets ont pour but de sécuriser la connexion configurée par les paquets "Initial".
- "Payload" : Une fois la connexion établie et sécurisée, ces paquets contiennent les données échangées.

QUIC est utilisé dans notre cas majoritairement pour des services Google.

E. UDP

Nous avons trouvé très peu d'utilisation du protocole UDP (environ 1 % d'utilisation), à l'exception de quelques échanges entre deux adresses IPv6 qui n'apparaissent qu'au début de la capture et que nous n'avons pas pu identifier.

V. Application

Dans la partie TCP IV-C, nous avons vu que plusieurs conversations TCP s'ouvraient. L'utilité de certaines était simple à comprendre vu le nombre de bytes échangés, d'autres l'étaient moins. Nous pensons donc à plusieurs possibilités, grâce à une analyse avec l'outil `Inspect / Network` intégré dans notre navigateur.

A. Modification en ligne

L'une d'entre elle est la modification collaborative en ligne. Il est intéressant de voir comment fonctionne la synchronisation des fichiers lorsque l'on modifie ceux-ci depuis le site.
Lorsque nous ouvrons un fichier éditable (du style `.txt`), nous remarquons différents appels à des fichiers JS. Tout d'abord, un appel vers `create` est effectué, permettant la création de l'environnement de modification. Ensuite, environ toutes les 5 secondes, un appel à `sync` permet la synchronisation avec le serveur, et donc aussi, si le fichier a été modifié ailleurs, de mettre à jour les données en direct. De plus, à chaque modification (par exemple à chaque touche sur laquelle on appuie), un appel à `push` est effectué, et permet de upload notre modification sur le serveur. Voici un exemple où l'on ajoute 4 lettres: 6.

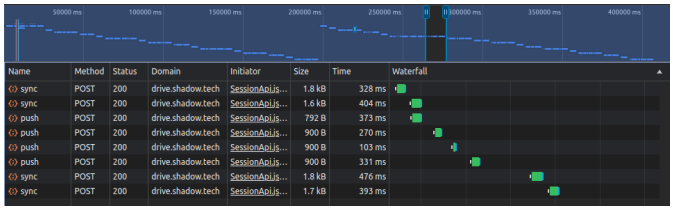

Fig. 6. Processus de modification et synchronisation d'un fichier

Ce processus est donc bien différent d'un upload ou download de fichiers, et s'effectue sans doute sur une des conversations TCP ouvertes. ShadowDrive inclut bien un système de modification en ligne collaboratif qui est fonctionnel.

B. Chargement des données

Nous pensons que d'autres conversations TCP sont utilisées simplement afin de charger les données du site, comme celles du chargement initial IV-B.

VI. Chiffrement et sécurité

En utilisant le site ssllabs [10], nous avons pu obtenir plus d'informations sur la sécurité de ShadowDrive, qui a reçu une note globale de A+ 7.

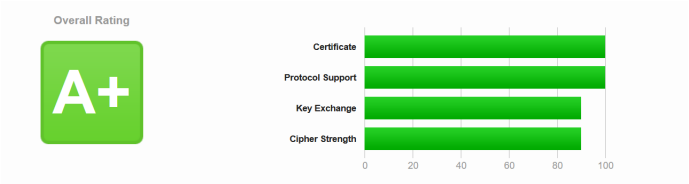

Fig. 7. Note globale de ShadowDrive

ShadowDrive supporte uniquement les 2 dernières versions de TLS, 1.2 et 1.3, qui sont reconnues pour être sécurisées. Cela implique que certaines anciennes versions de navigateur comme Safari 6 sur iOS 6.0.1 renvoient l'alerte : `handshake_failure`, qui indique que le client et le serveur ne peuvent pas établir une communication sécurisée. Dans ce cas-ci, cela est dû au fait que les suites de chiffrement proposées par Safari 6 ne sont soit pas sécurisées (ex: `TLS_ECDHE_ECDSA_WITH_RC4_128_SHA`) soit trop faibles (`TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA384`).

ShadowDrive utilise un certificat EV (Extended Validation) [11], c'est un certificat considéré comme ayant le plus haut niveau d'authentification pour un site Web. De plus, les utilisateurs peuvent voir le nom de l'entreprise dans la barre d'adresse ou bien une barre d'adresse verte pour être sûr

de l'authenticité du site. Pour obtenir un certificat de ce type, une CA (Certificate Authority) doit vérifier l'identité du propriétaire et son statut opérationnel en contrôlant le nom de domaine et le serveur d'hébergement. Pour ShadowDrive, le certificat a été émis par Sectigo RSA Extended Validation Secure Server CA et est valide du 30 octobre 2023 au 29 octobre 2024.

Pour la clé de cryptage utilisée, c'est une clé RSA de 2048 bits, elle offre un niveau de sécurité élevé et est actuellement considérée comme la norme pour les certificats SSL/TLS. `SHA256withRSA` nous indique que les données sont hachées à l'aide de l'algorithme `SHA-256`, qui est ensuite signé par la clé privée RSA de Sectigo. Cet algorithme permet de réduire les données à 256 bits indépendamment de la taille initiale des données.

Donc son certificat EV, sa clé RSA 2048 bits et l'algorithme SHA256withRSA garantissent la sécurité et l'authenticité de ShadowDrive.

VII. Conclusion

Ce rapport nous a permis de découvrir de nombreuses choses intéressantes. Nous en avons appris beaucoup sur les différents fonctionnements réseaux utilisés, ainsi que leurs utilités et leurs utilisations, appliquées à ShadowDrive dans ce cas. Cette analyse nous a aussi permis de prendre du recul sur l'outil que nous utilisons si souvent, sans pour autant connaître l'ampleur de celui-ci, ainsi que les précautions à prendre dans certaines situations.

References

[1] **M. Davan-Soulas, "Shadow Drive : un nouveau service de stockage en ligne français et ambitieux´´, 18/05/2022, consulté le 27/03/2024**
https://www.frandroid.com/produits-android/smartphone/1338483_shadow-drive-un-nouveau-service-de-stockage-en-ligne-francais-et-ambitieux

[2] **Multiple auteurs, "A successful student project - analyzing popular websites", 05/11/2019, consulté le 27/03/2024**
https://blog.computer-networking.info/project/

[3] **Multiple Auteurs, "Tunnellisation IPv6", 24/03/2023, consulté de 27/03/2024**
https://www.ibm.com/docs/fr/aix/7.3?topic=6-ipv6-tunneling

[4] **Baeldung (Multiple Auteurs), "Meaning of *** in the output of traceroute", 18/03/2024, consulté le 27/03/2024**
https://www.baeldung.com/linux/traceroute-three-starsµ

[5] **barryhunter, Forum Google, 12/01/2022, consulté le 27/03/2024**
https://support.google.com/webmasters/thread/145275615?hl=en& msgid=145329227

[6] **Detlef M., Forum Google, 28/04/2023, consulté le 27/03/2024**
https://support.google.com/pixelphone/thread/213101183?hl=en& msgid=213139418

[7] **SoftwareKeep, "What Is gstatic.com? Everything You Need To Know", consulté le 28/04/2023**
https://softwarekeep.com/nl-be/blogs/what-is/what-is-gstatic-com

[8] **Fastly, "Google Selects Fastly Oblivious HTTP Relay for Privacy Sandbox Initiative to Enhance Online Privacy for Billions of Chrome Users", consulté le 28/03/2024**
https://www.fastly.com/press/press-releases/google-selects-Fastly-Oblivious-HTTP-Relay-for-Privacy-Sandbox

[9] **Gregg Keizer, "Google launches leaked-password checker, will bake it into Chrome in December", 04/10/2019, consulté le 29/03/2024**
https://www.computerworld.com/article/3444237/google-launches-leaked-password-checker-will-bake-it-into-chrome-in-december.html

[10] **ssllabs, consulté le 01/04/2024**
https://www.ssllabs.com/ssltest/analyze.html?d=drive.shadow.tech&s=46.105.132.157

[11] **Wikipédia, "Extended Validation Certificate", dernière révision le 03/04/2024, consulté le 01/04/2024**
https:https://en.wikipedia.org/wiki/Extended_Validation_Certificate

[12] **www.ionos.fr, "QUIC : qu'est-ce qui se cache derrière le protocole expérimental de Google ¿', 01/03/2023, consulté le 02/04/2024**
https://www.ionos.fr/digitalguide/hebergement/aspects-techniques/quic/

[13] **Wikipedia, "Usage share of web browsers", dernière révision le 16/03/2024, consulté le 03/04/2024**
https://en.wikipedia.org/wiki/Usage_share_of_web_browsers