Méthodologie pour la caractérisation des relais d'attaques par déni de service distribué

Camille Moriot*, François Lesueur*, Nicolas Stouls*, Fabrice Valois*, Marie-Pierre Escudié†
*Univ Lyon, INSA Lyon, Inria, CITI, EA3720, 69621 Villeurbanne, France
†Univ Lyon, INSA Lyon, Institut Gaston Berger, 69621 Villeurbanne, France

***Résumé*—Les attaques par Déni de Service Distribué sont une menace constante sur l'écosystème Internet. Les solutions de détection et de mitigations sont de plus en plus performantes. Cependant, la majorité des solutions se concentrent au niveau de la victime, c'est à dire au bout de la chaîne de l'attaque. Nous proposons de mener un travail de métrologie de l'écosystème des DDoS et nous ciblons principalement les relais d'attaques dans le but d'identifier un point du réseau au plus près des sources qui permettrait d'affecter significativement l'ampleur des attaques. Pour cela, nous proposons une méthodologie pour caractériser les structures organisationnelles exploitées comme relais.**

I. Introduction

Depuis leur apparition, les attaques par déni de service distribué (DDoS) perturbent régulièrement Internet. Elles évoluent, conjointement et en réponse, avec les solutions proposées, et reposent sur l'exploitation de failles ou de fonctionnalités de systèmes divers. Le déni de service (DoS) peut être généré par deux mécanismes différents : l'afflux d'un volume de trafic qui sature le système victime ou l'épuisement des ressources protocolaires, logicielles ou architecturales de ce système. Afin d'amplifier le volume généré, l'attaquant coordonne des systèmes dits « relais » créant ainsi un DDoS (Figure 1a).

La grande majorité des solutions proposées au travers de la littérature aujourd'hui se localisent au niveau du réseau de la victime de l'attaque, essentiellement puisqu'elle peut intervenir sur son matériel ou en amont de son matériel via un service. Une autre approche, moins utilisée, au niveau des relais, a l'intérêt non négligeable de limiter la propagation de trafic malintentionné dans le réseau. Cependant, afin de pouvoir déployer des solutions efficaces au niveau des sources/relais DDoS, il est nécessaire de pouvoir les identifier et de les caractériser précisément.

L'étude des relais d'attaques DDoS est une composante essentielle pour comprendre tout l'écosystème qui en résulte. Des travaux ont apporté certaines clés de lecture, que nous présentons dans la section II. Nous voulons approfondir ces connaissances en apportant des informations quant à la nature des entités qui sont utilisées par les relais dans la section III. En suivant la méthodologie présentée en section IV, nous souhaitons établir des recommandations au niveau de la sécurité ainsi qu'identifier les points dans le réseau où le déploiement des solutions permettraient de réduire considérablement les effets des attaques DDoS.

II. État de l'art de l'étude des relais

En étudiant les attaques DDoS, les structure des réseaux d'attaques, les manières dont les attaquants s'emparent des systèmes relais ou encore la nature même du système sont des composantes connues aujourd'hui. Il est possible de remonter jusqu'aux systèmes relais en se basant par exemple sur l'adresse IP source du trafic DoS au niveau de la victime.

Les attaques DDoS reposent principalement sur des réseaux de machines (*les botnets*), contrôlés par l'attaquant via un moyen de communication. L'étude de ces réseaux est un point clé [1] et a permis le démantèlement de certains par la suite. L'étude de *botnets* comme celui employé lors de l'attaque Mirai en 2016 [2], permet de mettre en avant les mécanismes d'infections des relais, ici des objets IoT.

D'autres attaques reposent sur un mécanisme d'amplification. Dans ce cas, les attaques utilisent l'usurpation d'adresse IP, rendu possible par le mode non-connecté du protocole UDP, et l'attaquant fait réfléchir un trafic plus abondant vers la victime. C'est le cas par exemple des attaques par amplification DNS ou NTP [3].

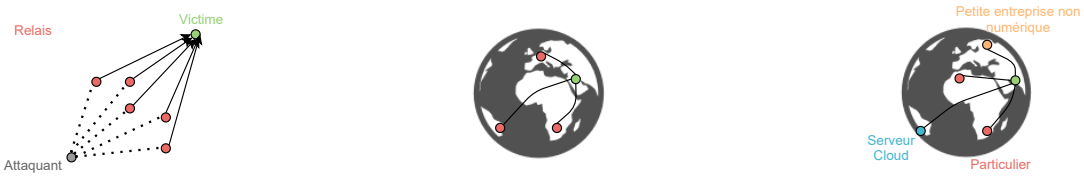
Enfin, aujourd'hui, des serveurs appartenant à des sociétés proposant des services d'hébergements, servent aussi de relais [4]. La qualité du débit, la possibilité de location de machines à l'heure, ainsi que l'accessibilité à ces machines sont des qualités attractives pour les attaquants.

En s'appuyant sur les adresses IPs sources du trafic DDoS reçu, l'origine géographique des relais d'attaques est aussi une composante largement traitée dans la littérature. Une analyse géographique des serveurs utilisés dans les attaques DDoS par amplifications a été mené dans [5].

D'autres travaux, s'appuyant aussi sur les adresses IPs, s'interessent à la nature organisationnelle des systèmes. Ainsi, dans [6], les auteurs ont proposé une méthode pour identifier le trafic résidentiel en se basant sur le contenu des réponses RDAP pour une adresse IP, qui leur permet de différentier le trafic résidentiel du reste du trafic.

III. Qualifier les relais d'attaques DDoS

Nous souhaitons donner un nouvel angle à l'analyse des relais d'attaque en examinant le type de structure à



(a) Attaque DDoS (b) Géolocalisation des relais (c) Caractérisation des relais

FIGURE 1: Connaissances des attaques DDoS

laquelle ils appartiennent. Actuellement, les analyses se concentrent sur la géolocalisation des relais (Figure 1b) tandis que nous nous intéressons à les caractériser plus finement (Figure 1c). Pour ce faire, nous avons choisi de nous focaliser sur deux axes : organisationnel et infrastructurel. L'objectif est de mieux comprendre les organisations qui sont derrière ces relais afin d'identifier comment accompagner leurs administrateurs pour lutter contre les DDoS.

Le premier axe a pour objectif de spécifier les structures (légales/organisationnelles) qui administrent les systèmes relais. Comprendre leurs structures nous permettra d'identifier les failles responsables de la présence de ces relais. Nous allons nous concentrer sur les étiquettes suivantes : le type d'organisation (entreprise, gouvernement, associations, particuliers), les activités de cette organisation, ainsi que le nombre de personnes faisant partie de la structure. Ces trois étiquettes permettraient de proposer une lecture qualifiant les moyens, ainsi que les connaissances qui peuvent être mises en place pour limiter la présence de ces relais. Ces informations sont une contribution nécessaire pour implémenter des solutions en adéquation avec la réalité du terrain.

Le second axe permettra d'évaluer le type d'infrastructure technique qui héberge les relais (taille, répartition géographique, services). Identifier le type d'infrastructure permettra de localiser précisément un point envisageable pour déployer une solution. De plus, ces étiquettes viendront enrichir une démarche d'analyse des vecteurs d'attaques et permettront d'établir les points techniques communs entre les réseaux hébergeant les relais.

IV. Méthodologie

Afin de réaliser ce travail de métrologie pointu, nous devons baser nos analyses sur des captures réelles d'attaques récentes. Celles-ci doivent être prélevées à des points stratégiques du réseau : au niveau du réseau des victimes pour garantir la pluralité des sources et au niveau d'intermédiaires régulièrement visés par ces types d'attaques. Nous envisageons actuellement deux natures de points : au niveau des réseaux universitaires et au niveau des entreprises régulièrement visées par ces attaques.

Notre contribution proposera une qualification du type d'IP en se basant sur un recoupement de différents services tels que Wikidata, Kompass ou RDAP. Les proportions de chacune des étiquettes seront étudiées en fonction du volume généré par chaque nature de relais.

Certaines difficultés sont à prendre en compte dans la méthodologie. Tout d'abord, l'adresse IP étant une donnée personelle, il faut préserver l'identité des relais. Pour cela, les adresses IP des captures seront converties vers les adresses des sous-réseaux. De plus, l'affectation de l'adresse IP peut avoir changé entre le moment de la trace et celui de l'analyse. Il nous faut donc mémoriser l'état du réseau au plus proche de la date de la trace. Enfin, il sera nécessaire de prendre en compte le taux d'usurpation des adresses IP. Notamment, nous devrons mettre en évidence le taux d'adresses sources dont nous ne sommes pas en mesure d'affirmer si elles sont licites ou usurpées.

V. Conclusion et perspectives

Nous avons proposé une méthodologie, basée sur les adresses IP, pour qualifier les structures techniques mais aussi socio-organisationnelles hébergeant des relais d'attaques DDoS. C'est un travail de métrologie de l'écosystème qui se voudra au plus proche de la réalité du terrain, donc une méthodologie que nous allons appliquer à des traces d'attaques réelles capturées aux points stratégiques identifiés.

Références

[1] A. Wang, W. Chang, S. Chen, and A. Mohaisen, "Delving Into Internet DDoS Attacks by Botnets : Characterization and Analysis," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2843–2855, 2018.

[2] R. Hallman, J. Bryan, G. Palavicini, J. Divita, and J. Romero-Mariona, "IoDDoS-the internet of distributed denial of sevice attacks," in *2nd international conference on internet of things, big data and security.*, 2017, pp. 47–58.

[3] C. Rossow, "Amplification Hell : Revisiting Network Protocols for DDoS Abuse," in *21st Annual Network and Distributed System Security Symposium, NDSS,USA*, 2014.

[4] Z. He, T. Zhang, and R. B. Lee, "Machine Learning Based DDoS Attack Detection from Source Side in Cloud," in *4th IEEE International Conference on Cyber Security and Cloud Computing, CSCloud 2017, USA*, 2017, pp. 114–120.

[5] J. J. Santanna, R. van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Z. Granville, and A. Pras, "Booters - An analysis of DDoS-as-a-service attacks," in *IFIP/IEEE International Symposium on Integrated Network Management, IM 2015, Canada*, 2015, pp. 243–251.

[6] X. Mi, X. Feng, X. Liao, B. Liu, X. Wang, F. Qian, Z. Li, S. A. Alrwais, L. Sun, and Y. Liu, "Resident evil : Understanding residential IP proxy as a dark service," in *IEEE Symposium on Security and Privacy, SP, USA*, 2019, pp. 1185–1201.