



HAL
open science

Zeph & Iris cartographient l'internet

Matthieu Gouel, Kevin Vermeulen, Maxime Mouchet, Justin Rohrer, Olivier Fourmaux, Timur Friedman

► **To cite this version:**

Matthieu Gouel, Kevin Vermeulen, Maxime Mouchet, Justin Rohrer, Olivier Fourmaux, et al.. Zeph & Iris cartographient l'internet. CORES 2022 – 7ème Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication, May 2022, Saint-Rémy-Lès-Chevreuse, France. hal-03656974

HAL Id: hal-03656974

<https://hal.science/hal-03656974>

Submitted on 2 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Zeph & Iris cartographient l'internet[†]

Matthieu Gouel^{1,2}, Kevin Vermeulen³, Maxime Mouchet^{1,2}, Justin P. Rohrer⁴,
Olivier Fourmaux¹ et Timur Friedman^{1,2}

¹Sorbonne Université, CNRS, LIP6, 4 place Jussieu, 75005 Paris

²LINCS, 23 avenue d'Italie, 75013 Paris

³LAAS-CNRS, 7 avenue du Colonel Roche, 31031 Toulouse

⁴TaNCAD Lab, Naval Postgraduate School, 1 University Circle, Monterey, CA 93943, États-Unis

Une cartographie plus complète de la topologie d'internet permet de mieux comprendre des propriétés importantes d'internet, comme sa résilience et sa performance. Nous présentons deux contributions pour fournir une topologie plus complète à la communauté : (1) Zeph, un algorithme d'apprentissage par renforcement qui alloue les préfixes à sonder pour chaque point de mesures afin d'optimiser la découverte de la topologie IP d'internet. (2) Iris, une plateforme de mesure open-source capable de conduire des mesures de topologie exhaustives sur l'ensemble d'internet.

Mots-clefs : mesures actives d'internet; topologie d'internet; traceroute

1 Introduction

La cartographie de la topologie d'internet est un défi pour la communauté scientifique depuis plus de deux décennies. La topologie d'internet peut avoir plusieurs niveaux de granularité : le niveau des interconnexions entre les organisations qui composent le réseau, le niveau des interconnexions entre les routeurs, et le niveau des chemins IP des interfaces des routeurs. Ce dernier niveau de granularité peut être partiellement découvert en effectuant des mesures de type *traceroute* [Jac88] depuis des points de mesures vers des destinations d'internet (chaque préfixe /24 routé d'internet). On peut ainsi révéler des nœuds du réseau (adresses IP) et des liens entre ces nœuds (deux adresses IP consécutives d'un chemin).

Le système de l'état de l'art Ark [Hyu06] utilise une infrastructure qui ne lui permet pas de lancer des mesures vers toutes les destinations d'internet depuis toutes leurs sources, car le nombre de paquets par seconde que les points de mesures peuvent envoyer est limité. En conséquence, Ark distribue l'espace des destinations aléatoirement entre chaque point de mesures à chaque cycle de mesure et recommence au cycle suivant avec une nouvelle distribution aléatoire.

Dans cet article, nous présentons Zeph et Iris. Zeph est un algorithme d'apprentissage par renforcement qui sélectionne les préfixes à sonder pour chaque point de mesures de manière à maximiser la découverte de nœuds et de liens. Zeph repose sur les observations suivantes : (1) Les routes d'internet entre une source et une destination sont globalement stables pour des jours entiers [GKF⁺20] (2) La structure des chemins d'internet offre une redondance dans ces chemins [DRFC06], et beaucoup de paquets découvrent les mêmes nœuds/liens. La stabilité des routes d'internet à l'échelle de la journée permet de s'appuyer sur des mesures passées pour sélectionner quels préfixes doivent être attribuées dans le futur à quels points de mesures afin de maximiser les découvertes. De plus, la redondance des chemins permet d'envisager de ne sonder qu'une partie des préfixes en se concentrant sur ceux qui permettent de découvrir des nœuds et des liens uniquement visibles depuis un seul point de mesures particulier. Ainsi, Zeph apprend progressivement quelles sont les destinations à sonder de préférence pour chaque point de mesures, et l'apprentissage par renforcement aux dynamiques de la topologie d'internet au cours du temps. Iris est le système de mesures à l'échelle d'internet open source qui supporte l'implémentation de Zeph.

[†] « Zeph & Iris Map the Internet - A resilient reinforcement learning approach to distributed IP route tracing » est paru dans la revue ACM SIGCOMM Computer Communication Review de Janvier 2022.

Grâce à cette nouvelle approche, Zeph est capable de découvrir 3,3 millions de nœuds et 19.9 millions de liens en 15 heures, soit 3 fois plus de nœuds et 10 fois plus de liens que l'état de l'art.

Les codes sources d'Iris et Zeph sont rendus publics sous licence MIT, et nous fournissons à intervalle régulier des données sur la topologie IP d'internet[‡].

2 L'orchestrateur Zeph

Comme présenté en introduction, l'objectif de Zeph est d'optimiser le nombre de liens de la topologie IP d'internet découverts au cours d'un cycle de mesures. À chaque cycle, Zeph prend en paramètres un ensemble de points de mesures, nommés agents, et le budget de chaque agent, i.e., le nombre de destinations qu'il peut sonder en un cycle avec l'outil de mesure type `traceroute`[§] lors d'un cycle.

L'ensemble des destinations possibles est l'ensemble des préfixes /24 routables extraites des tables BGP obtenues via `RouteViews`. À chaque cycle, Zeph alloue un ensemble de destinations que chaque agent doit sonder, puis chaque agent sonde ces destinations. Pour sélectionner ces paires (agent, destination), Zeph procède par apprentissage par renforcement. Les paires provenant de *l'exploitation* sont calculées d'après les découvertes du cycle précédent, en s'appuyant sur la nature relativement statique et redondante des routes d'internet [GKF⁺20], et sont choisies en fonction des découvertes faites au cycle précédent. Les paires provenant de *l'exploration* sont sélectionnées aléatoirement afin de découvrir d'éventuelles routes apportant de meilleures découvertes, par exemple dues à des changements de routages dans internet [GKF⁺20]. L'ensemble des paires (agent, destination) du premier cycle provient exclusivement de l'exploration, car il n'y a pas de mesure précédente sur laquelle s'appuyer pour effectuer l'exploitation. La part minimum du budget que chaque agent dédie à l'exploration est un paramètre de Zeph, appelé ϵ dans la littérature, que nous fixons à 0.1 (10%). Nous évaluons ce choix dans la Sec. 4. Les sections suivantes détaillent comment s'effectuent l'exploitation et l'exploration.

2.1 Exploitation

Zeph s'appuie sur la stabilité des routes d'internet [GKF⁺20] pour allouer des destinations aux différents agents : si un agent sonde les mêmes destinations lors de deux cycles consécutifs, on s'attend à ce que chaque mesure redécouvre les mêmes liens, modulo des dynamiques d'internet comme les changements de routage ou les pannes. Pour sélectionner les paires (agent, destination) qui doivent être mesurées de nouveau de manière à maximiser les découvertes, Zeph formalise le problème sous forme de couverture par ensemble : l'ensemble à couvrir est l'ensemble des liens formés par l'union des liens découverts par tous les chemins IP mesurés par chaque agent. Un sous-ensemble à sélectionner est un ensemble de liens découvert par un `traceroute`, i.e., une paire (agent, destination). L'objectif est donc de sélectionner le minimum de paires (agent, destination) pour couvrir l'ensemble des liens découverts. Ce problème étant NP complet, Zeph utilise une variante de l'algorithme glouton pour les grands jeux de données, appelé `Disk-Friendly Greedy algorithm (DFG)` [CKW10]. La sortie de DFG est donc un ensemble de paires (agent, destination). Chaque agent se voit allouer les destinations qui lui sont associées dans les limites de son budget d'exploitation (90% de son budget total).

2.2 Exploration

À la fin de l'exploitation, chaque agent se voit attribuer un ensemble de destinations à mesurer. Pour chaque agent, Zeph attribue en plus au moins 10% du budget de l'agent (ou plus si l'exploitation n'a pas rempli les 90% disponibles) à de nouvelles destinations sélectionnées aléatoirement dans l'ensemble des destinations disponibles pour cet agent et qui n'ont pas été sondées au cycle précédent.

‡. Les codes sources d'Iris et Zeph sont disponibles sur GitHub : <https://github.com/dioptra-io>; les données sont disponibles sur notre site internet : <https://iris.dioptra.io>.

§. Zeph supporte plusieurs outils, comme `Yarrp` [Bev16] ou `Diamond-Miner` [VRB⁺20], qui permet de tracer les chemins à répartition de charge.

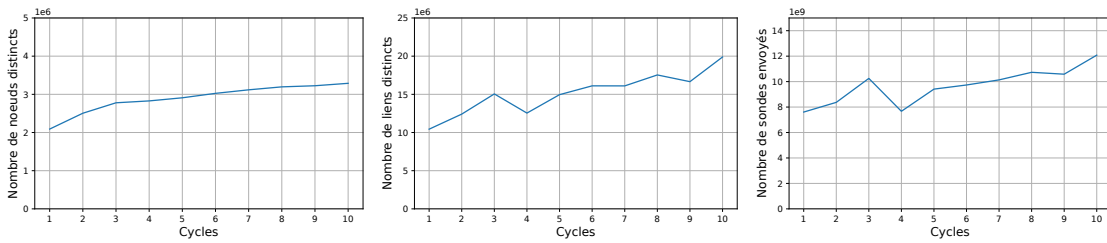


FIGURE 1: Nombre de nœuds et de liens découverts, ainsi que le nombre de sondes envoyées, pour chaque cycle de Zeph utilisant Diamond-Miner. Mesure effectuée entre le 21 janvier et le 28 janvier 2022 depuis 5 nœuds EdgeNet localisés dans les régions Google Cloud suivantes : asia-east2-a, asia-northeast1-a, asia-southeast1-a, europe-west6-a et southamerica-east1-a. Au minimum 10% du budget de mesure est alloué à l’exploration, la probabilité d’échec de Diamond-Miner est réglée à 5% et 100 000 sondes sont envoyées par seconde. Les versions suivantes des outils sont utilisées : Iris v1.0.0, Zeph v1.0.0 et Diamond-Miner v0.6.1.

3 La plateforme de mesure Iris

Afin de conduire des mesures régulières de traceroutes multi-chemins à l’échelle d’internet, nous avons développé la plateforme de mesure Iris . Cette plateforme a non seulement pour but de supporter nos propres recherches, mais aussi de fournir des données régulières de la topologie d’internet à la communauté scientifique. Il est également possible pour un utilisateur de conduire ses propres mesures via notre plateforme ou bien de déployer sa propre plateforme sur son infrastructure, car le code complet d’Iris et sa documentation est open-source.

À l’heure actuelle, Iris est capable de supporter des mesures de type traceroute avec YARRP [Bev16], Diamond-Miner [VRB⁺20], ainsi que des mesures de type Ping. Ces mesures peuvent se faire via les protocoles IPv4 ou IPv6. Iris se pilote via son API REST. Grâce à celle-ci, nous pouvons facilement intégrer des outils d’orchestration de mesures tels que Zeph .

4 Évaluation

Les résultats principaux de Zeph sont :

Zeph, utilisé avec Iris et Diamond-Miner fournit les données les plus complètes de la topologie IPv4 d’internet. Nous avons comparé notre système complet avec l’état de l’art des systèmes de mesure de la topologie d’internet Ark, opéré par CAIDA. Zeph+Iris est capable de découvrir en 15 heures 20 millions de liens en sondant depuis seulement 5 points de mesures à 100 000 paquets par seconde. En comparaison, Ark est capable d’atteindre le même niveau de découverte en 45 jours avec plus de 100 points de mesures répartis dans le monde sondant à 100 paquets par seconde. Ce temps de mesure largement plus court est important, car cela permet de produire des topologies d’internet plus exactes, car moins sensibles aux dynamiques. La Fig. 1 montre l’évolution sur 10 cycles des découvertes de nœuds et de liens découverts ainsi que le nombre de sondes. On peut noter qu’un agent s’est déconnecté durant le cycle 4 mais Zeph s’est adapté et a repris une évolution croissante dans les cycles suivants. On peut voir que l’apprentissage par renforcement fonctionne : au cycle 1 Zeph a découvert 2,1M nœuds et 10,4M liens tandis qu’au cycle 10 celui-ci a découvert 3,3M nœuds (+57.3%) et 19,8M liens (+90,5%). Néanmoins, on peut noter une augmentation de 58,7% du nombre de sondes utilisé entre le cycle 1 et le cycle 10.

Zeph permet de réduire le nombre de sondes envoyées. En comparant Zeph avec différents budgets de préfixes par agent (Fig. 2 a), nous avons évalué que Zeph permet de découvrir quasiment le même nombre de nœuds et de liens avec seulement 50% des préfixes sondés, permettant d’économiser ce même pourcentage de sondes (742 millions de sondes contre 1,485 milliard de sondes).

Zeph est meilleur qu’une allocation de préfixes aléatoire. Nous avons mis au point des expériences permettant de comparer Zeph avec d’autres stratégies d’allocation de préfixes tout en limitant les facteurs de confusion liés à la vitesse de sondage, le nombre de points de mesures, leur puissance, leur connectivité

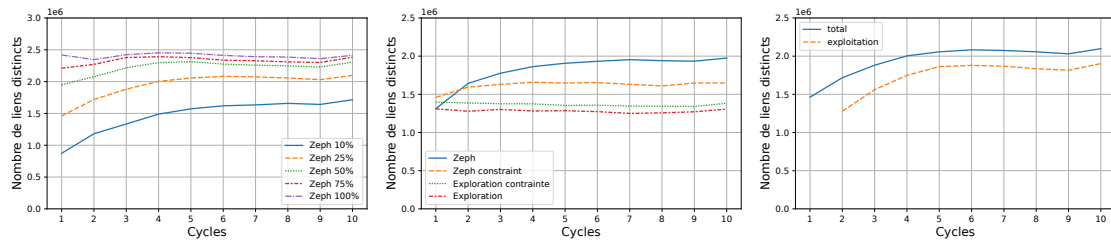


FIGURE 2: Évaluation de Zeph. À gauche (a), comparaison de mesures Zeph sur 10 cycles avec différents budgets par agent. Au centre (b), comparaison de Zeph avec d'autres stratégies d'allocation de préfixes. À droite (c), découvertes totales de Zeph en comparaison avec les découvertes de la partie exploitation du budget des agents uniquement.

et leur position dans le réseau. En particulier, nous voulions comparer Zeph avec une distribution totalement aléatoire, ainsi qu'une distribution avec la contrainte de sonder un même préfixe depuis un seul point de mesures comme le fait Ark. La Fig. 2 (b) compare 4 stratégies d'allocation de préfixes : (1) Zeph (2) Zeph contraint (3) une allocation aléatoire (4) une allocation aléatoire contrainte. On peut constater que la stratégie de Zeph dépasse rapidement les autres stratégies à partir du cycle 3.

L'exploitation de Zeph rapporte l'essentiel des découvertes. La Fig. 2 (c) montre le nombre total de liens découverts et le nombre de liens découverts seulement avec les préfixes de l'exploitation. On peut voir que la grande majorité des découvertes (90%) vient de la partie exploitation bien que la partie exploration permette d'augmenter le nombre de découvertes cycles après cycles. Durant cette évaluation, nous nous sommes également rendu compte que Zeph utilisait seulement environ 10% du budget pour l'exploitation au lieu des 90% maximum paramétrés par ϵ . Une grande partie des découvertes est donc faite avec très peu de préfixes (10%) et le reste du budget est utilisé à trouver des préfixes plus optimaux.

5 Conclusion

Zeph est un nouvel algorithme capable d'optimiser la découverte de liens de la topologie d'internet. Utilisé avec Iris, notre nouvelle plate-forme open-source de mesures sur internet, nous sommes capables de découvrir 3 fois plus de nœuds et 10 fois plus de liens pour le même nombre de chemins sondés, et ce en un temps plus court. Nous produisons des données régulières de la topologie d'internet qui peuvent être téléchargées par la communauté scientifique via notre site web.

Références

- [Bev16] Robert Beverly. Yarrp'ing the internet : Randomized high-speed active topology discovery. In *ACM IMC*, 2016.
- [CKW10] Graham Cormode, Howard Karloff, and Anthony Wirth. Set cover algorithms for very large datasets. In *ACM CIKM*, 2010.
- [DRFC06] B. Donnet, P. Raoul, T. Friedman, and M. Crovella. Deployment of an algorithm for large-scale topology discovery. *IEEE JSAC*, 2006.
- [GKF⁺20] V. Giotsas, T. Koch, E. Fazzion, I. Cunha, M. Calder, H. V. Madhyastha, and E. Katz-Bassett. Reduce, reuse, recycle : Repurposing existing measurements to identify stale traceroutes. In *ACM IMC*, 2020.
- [Hyu06] Y. Hyun. The Archipelago Measurement Infrastructure. In *7th CAIDA-WIDE Workshop*, 2006.
- [Jac88] Van Jacobson. 4BSD routing diagnostic tool available for ftp. Email 8812201313.AA03127@helios.ee.lbl.gov to the IETF and end2end-interest e-mail lists, 1988.
- [VRB⁺20] K. Vermeulen, J. P Rohrer, R. Beverly, O. Fourmaux, and T. Friedman. Diamond-Miner : Comprehensive Discovery of the Internet's Topology Diamonds. In *NSDI*, 2020.