Université de Strasbourg

École doctorale MSII

Détection des éléments défaillants dans l'Internet : mensonges inter-domaines, détours intra-domaine et anomalies liées aux points d'échanges

THÈSE

pour obtenir le grade de

Doctorat de l'Université de Strasbourg

(mention informatique)

28 novembre 2019

présentée par

Julián Martín DEL FIORE

Composition du jury

Directeur de thèse:	Cristel PELSSER, ICube, Université de Strasbourg, France
Co-encadrant:	Pascal MERINDOL, ICube, Université de Strasbourg, France
Rapporteurs:	Benoit DONNET, Department of Electrical Engineering and Computer Science Montefiore Institute, University of Liège, Belgium
	Guillaume URVOY-KELLER, I3S Research Laboratory, University of Nice Sophia Antipolis, France
Examinateurs:	Clémence MAGNIEN, LIP6 laboratory, Sorbonne Université, France

Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie
ICUBE — UMR 7357

Sommaire de Thèse

L'**objectif** de cette thèse est de **détecter des éléments défaillants de l'Internet**, c'est-à-dire des composants défectueux, des réseaux confrontés à des limitations et des réseaux privilégiant leurs propres revenus au détriment d'une meilleure performance de leurs services Internet.

L'**Internet** peut apparaître comme un **système infaillible**, qui n'échoue jamais, de par sa dimension distribuée, cependant ce n'est pas le cas. Internet est simplement une interconnexion de réseaux indépendants, connue sous le nom de **Systèmes autonomes (ASes)**. Étant donné que les ASes sont construits sur base d'un **hardware** et d'un **software** potentiellement défectueux, et que les **opérateurs de réseaux**, des agents **humains** par nature sujet aux erreurs, gèrent eux-mêmes les AS, alors Internet exhibe certains **limites**. Certaines **décisions arbitraires** prisent par les opérateurs ne sont pas toujours optimales. De plus, pour un même problème, différentes personnes peuvent considérer des contraintes incohérentes comme les plus pertinentes, et donc proposer des **solutions divergentes**, contribuant à une couche supplémentaire de défaillances dans le comportement d'Internet. D'un point de vue économique, le fait que certaines entreprises soient sujettes à l'**avarice** transforme de facto l'Internet en un **écosystème axé sur le profit**. De plus, les **systèmes construits par l'homme** ne sont jamais parfaits : ils ont tendance à inclure des fonctionnalités qui peuvent **échouer** et exigent de la **maintenance**. Après un temps donné, le système peut devenir **obsolète** et exiger des **mises à jours**. Tous ces facteurs amène l'Internet à inclure de nombreux **éléments défaillants**, c'est-à-dire des **composants défectueux**, des réseaux confrontés à des **limitations** et même **réseaux égoïstes** donnant la priorité à leurs propres revenus plutôt qu'à de meilleures performances de l'Internet. Il s'agit de ce que nous cherchons à **détecter** dans cette thèse. Ce défi est **difficile** à relever : les phénomènes que nous voulons identifier et analyser peuvent se produire n'importe où, ou seulement dans certains AS, alors qu'Internet compte aujourd'hui environ 70K ASes (en novembre 2020).

Le premier composant de l'Internet que nous étudions dans cette thèse sont les **points d'échange Internet (IXP)**, les installations d'interconnexion couramment utilisées par les ASes. La structure d'Internet, c'est-à-dire la manière dont les AS établissent des connexions entre eux, a été largement modifiée par l'irruption des IXP dans les années 2000. Les IXP permettent aux AS d'établir des connexions à une plus grande échelle et de produire des **économies monétaires**. Cependant,

la popularité des IXP varie selon les régions. Dans certains cas, les pays peuvent avoir **IXP déficients**, c'est-à-dire aucun IXP, ou un IXP qui n'a pas réussi à attirer suffisamment de membres. En général, les IXP sont connus pour avoir eu beaucoup de succès en Europe, où les pays et surtout les distances sont relativement petites par rapport aux autres régions. Ce succès ne s'est pas illustré dans la même mesure en Amérique du Nord. Des études récentes se sont aussi concentrées sur le rôle des IXP dans l'écosystème Internet africain. Dans d'autres régions, en revanche, on en sait peu sur les IXP et même sur l'Internet dans son ensemble. En particulier, compte tenu des progrès de l'Internet dans la région et de la taille de la population, **Amérique latine (LatAm)** représente une région intéressante à étudier. Cela nous amène à notre première série de questions de recherche.

Question de Recherche

- **Combien d'IXP sont déployés à travers LatAm ?**
- **Les IXP de LatAm contribuent-ils à l'écosystème AS de manière nationale, régionale ou internationale ?**
- **Tous les IXP de LatAm ont-ils réussi à proliférer ou y a-t-il des IXP déficitaires ? Si certains ont échoué, pourquoi ?**

Le déploiement des IXP en LatAm est notre première étude d'un élément potentiellement défaillant de l'Internet. En effet, les IXP sont une histoire de **succès ou échec** selon le pays considéré en Amérique latine. Alors que l'Argentine, le Brésil et le Chili comptent avec des IXP qui ont réussi à proliférer, d'autres pays ont des **IXP défaillants**. Notre étude caractérise non seulement l'Internet dans la région, mais explore également les raisons pour lesquelles certains IXP pourraient réussir tandis que d'autres non.

Le deuxième élément de l'Internet que nous examinons dans cette thèse est le protocole de routage inter-domaine utilisé à large échelle dans l'Internet : **border gateway protocol (BGP)**. BGP dicte comment les ASes échangent des informations d'accessibilité concernant les préfixes IP que chacun d'eux possède. A haut niveau, chaque AS annonce à ses AS voisins les préfixes qu'il possède dans BGP, et à leur tour, ceux-ci relaient le message à d'autres AS. Dans ce processus distribué, les messages de routage qui sont envoyés gardent une trace du chemin AS qui a été suivi, c'est-à-dire une liste des AS, du premier au dernier, qui ont annoncé le préfixe. Une analogie d'un AS annonçant un préfixe avec un chemin AS donné vers un AS voisin, est l'offre d'un **contrat**. Plus précisément, l'annonce BGP joue le rôle du contrat, et le service offert est que, pour atteindre un préfixe donné, les paquets circuleront en suivant l'ensemble ordonné d'AS exprimé dans le chemin AS. Pour cette raison, le chemin que suivent effectivement les paquets, le réel chemin de commutation, devrait **correspondre** aux chemins AS qui sont annoncés pour tous les préfixes. Lorsque les deux chemins **divergent**, nous dirons que les AS se livrent à des **mensonges BGP**, et cela génère une deuxième défaillance de l'Internet, exploitant la

naïveté du protocole BGP. Cela nous amène à une seconde série de questions de recherche.

Question de Recherche

- **Existe-t-il des mensonges BGP sur Internet ?**
- **Pouvons nous décrire comment un AS génère ses mensonges BGP ?**
- **Pouvons nous développer un cadre d'analyse pour détecter ces mensonges ?**

Comme BGP est au cœur de l'Internet, nous pouvons aisément affirmer que l'Internet est très dépendant de BGP, d'où l'importance critique de s'assurer que celui-ci fonctionne bien et que les chemins d'AS sont respectés en pratique. Cependant, si cela se produit réellement, il n'est pas facile de le vérifier. Par conséquent, la **confiance implicite** selon laquelle les AS annoncent les chemins qu'ils utilisent pour la transmission de paquets peut être **naive**. Au-delà d'affecter éventuellement l'optimalité des chemins, les mensonges BGP peuvent avoir des **implications pour la sécurité**, c'est-à-dire se rapporter à un comportement intéressé où les opérateurs de réseaux **malveillants** ont des motivations éonomiques ou politiques pour détourner le trafic de ceux actuellement proposés. Par exemple, le trafic redirigé pourrait être livré à une partie malveillante tentant de recueillir des informations à partir des paquets collectés. En même temps, puisque les AS ne sont pas exempts d'erreurs, le même mensonge BGP pourrait résulter de **limitations techniques** et le responsable opérateur pas nécessairement malveillant, ni même au courant. Dans cette thése nous modélisons comment les AS peuvent introduire des mensonges BGP, et nous proposons une méthodologie pour les détecter. Cette tâche nécessite de relever un défi pratique considérable : les données collectées à partir des mesures et des outils nécessaires à une telle analyse sont **bruités**, par conséquent, les mensonges BGP peuvent être mal interprétés (comme du bruit, ou vice versa). Notre cadre d'analyse permet de **filtrer** le bruit affectant les mesures, et de tirer des conclusions représentatives.

Le dernier composant de l'Internet que nous analysons sont les **protocoles de routage intra-domaine (IGP)**, les protocoles que les AS utilisent à l'intérieur même de leurs réseaux. Les IGP se caractérisent par un schéma de transfert tel que le trafic traversant les AS emprunte les meilleurs chemins qui **minimisent** la distance, ou le coût, selon une métrique de leur choix, souvent additive et prenant en compte les délais d'émission et de propagation. Pourtant des **détours** peut se produire, c'est-à-dire que le trafic peut circuler à travers des routes de transfert qui divergent des meilleurs chemins attendus. Cela nous amène à un troisième ensemble de questions de recherche.

Question de Recherche

- **Est-il fréquent que des détours se produisent sur l'Internet ?**
- **Pouvons-nous modéliser formellement les causes produisant des détours de transfert ?**
- **Ces causes premières génèrent-elles toujours des détours de transfert, ou pourraient-elles modifier la transmission d'une autre manière ?**
- **Peut-on concevoir une méthodologie qui, sans nécessiter de connaissances privilégiées des réseaux analysés, est capable de détecter si des détours se produisent à l'intérieur de ceux-ci ?**

Le concept de détours se focalise sur les conséquences, c'est-à-dire sur la différence entre les chemins observés et les vrais meilleurs chemins IGP, mais ne dit rien sur la façon dont ceux-ci sont générés, c'est-à-dire sur les **origines** d'un tel phénomène, que nous aimerions étudier. A priori, les détours peuvent résulter des effets secondaires de **solutions de contournement de l'évolutivité**, où les AS tentent de prolonger la vie de leurs **anciens routeurs** en acceptant un **compromis** des performances de leur réseau. De plus, des **bogues** dans le logiciel des routeurs peuvent également créer des détours de transfert, et par conséquent les opérateurs de réseau peuvent **ignorer** que ceux-ci se produisent dans leurs AS. Contrairement au principe de base de la métrique IGP (tout est meilleur chemin), les détours de transfert **augmentent** la distance IGP requise pour traverser un AS et aboutissent sans doute à **gaspillage des ressources** à l'intérieur du réseau.

En particulier, dans cette thése nous décrivons les effets des détours de transfert et pourquoi **les détecter est complexe**. L' **outil** que nous développons pour aborder cet objectif est sophistiqué. En effet, c'est un défi particulièrement difficile à relever car les détours et les meilleurs chemins IGP sont simplement des routes de transfert utilisées à l'intérieur des AS, un agent externe ne peut pas facilement les distinguer (sans connaître la métrique et la topologie IGP utilisée à l'intérieur de l'AS analysé). De plus, il existe d'autres techniques, telles que **l'équilibrage de charge et l'ingénierie de trafic** qui, en général, introduisent également des chemins alternatifs à un unique meilleur chemin IGP. Contrairement aux détours, ceux-ci sont toujours considérés comme optimaux au sens du coût IGP (avec ECMP et les multi-chemins de coût égaux) ou pour les besoins spécifiques de l'AS qui les déploie, respectivement. En particulier, nous présentons un **modèle de commutation multi-niveau** permettant de décrire formellement les détours. Pour finir, nous décrivons également l'outil que nous avons conçu pour détecter les détours en filtrant les chemins issus de l'équilibrage de charge et de l'ingénierie de trafic.

Liste de publications

Journals

- **Julián M. Del Fiore**, Valerio Persico, Pascal Merindol, Cristel Pelsser and Antonio Pescapè. *The Art of Detecting Forwarding Detours*, submitted to IEEE Transactions on Network and Service Management (IEEE TNSM).

- Esteban Carisimo, **Julián M. Del Fiore**, Diego Dujovne, Cristel Pelsser, and J. Ignacio Alvarez-Hamelin. 2020. *A first look at the Latin American IXPs*, in SIGCOMM Comput. Commun. Rev. 50, 1 (January 2020), 18–24. DOI:https://doi.org/10.1145/3390251.3390255

- R. Koutsiamanis, G. Z. Papadopoulos, X. Fafoutis, **Julián M. Del Fiore**, P. Thubert and N. Montavont, *From Best Effort to Deterministic Packet Delivery for Wireless Industrial IoT Networks*, in IEEE Transactions on Industrial Informatics (IEEE TII), vol. 14, no. 10, pp. 4468-4480, Oct. 2018, doi: 10.1109/TII.2018.2856884.

Conferences

- Esteban Carisimo, **Julián M. Del Fiore**, J. Ignacio Alvarez-Hamelin, Cristel Pelsser, *LatAm's IPv6 Uptake From an IXP Perspective*, to be submitted.

- Esteban Carisimo, **Julián M. Del Fiore**, J. Ignacio Alvarez-Hamelin, Cristel Pelsser, *LatAm's IXPs in The Global Landscape*, submitted to Passive and Active Measurement (PAM) Conference 2021.

- **Julián M. Del Fiore**, Pascal Merindol, Valerio Persico, Cristel Pelsser and Antonio Pescapè. *Filtering the Noise to Reveal Inter-Domain Lies*, in 2019 Network Traffic Measurement and Analysis Conference (TMA), pages 17–24, 2019, IEEE.

- M. Kersalé, G. Z. Papadopoulos, **Julián M. Del Fiore**, P. Thubert, N. Montavont. *Towards deterministic industrial network*, in Rencontres Francophones sur la Conception de Protocoles,l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication (CoRes 2017), May 2017, Quiberon, France.

Posters

- **Julián M. Del Fiore**, Pascal Merindol, Valerio Persico, Cristel Pelsser and Antonio Pescapè. *Routing Inconsistencies at the FIB level*, in 2019 Network Traffic Measurement and Analysis Conference (TMA), IEEE.
- Esteban Carisimo, **Julián M. Del Fiore**, Diego Dujovne, Cristel Pelsser, J. Ignacio Alvarez-Hamelin, *Country-level influence of IXPs in Latin America*, in Latin American Student Workshop on Data Communication Networks (LANCOMM) 2019.
- **Julián M. Del Fiore**, Pascal Merindol, Valerio Persico, Cristel Pelsser and Antonio Pescapè. *A BGP-lying Tale: Stop Blamming the Mapping*, in 2018 Network Traffic Measurement and Analysis Conference (TMA), IEEE.