



**HAL**  
open science

# Distribution vidéo sur Internet : Des mesures P2P-TV à la cartographie d'Internet

Olivier Fourmaux

► **To cite this version:**

Olivier Fourmaux. Distribution vidéo sur Internet : Des mesures P2P-TV à la cartographie d'Internet. Réseaux et télécommunications [cs.NI]. Sorbonne Université, 2021. tel-03721310

**HAL Id: tel-03721310**

**<https://hal.science/tel-03721310>**

Submitted on 12 Jul 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.

**Mémoire d'habilitation à diriger des recherches  
spécialité informatique**

## **Distribution vidéo sur Internet**

**Des mesures P2P-TV à la cartographie d'Internet**

Olivier Fourmaux

Soutenance prévue le 9 décembre 2021  
version 1.1 (jury)

Devant le jury composé de :

Toufik Ahmed	Professeur, Bordeaux INP (rapporteur)
Anne Fladenmuller	Professeure, Sorbonne Université
Stefano Secci	Professeur, CNAM
Sébastien Tixeuil	Professeur, Sorbonne Université
Sandrine Vaton	Professeure, IMT Atlantique (rapporteuse)
Véronique Vèque	Professeure, Université Paris-Saclay (rapporteuse)



# Remerciements

Tout d'abord, je tiens à exprimer toute ma gratitude à Toufik Ahmed, Sandrine Vaton et Véronique Vêque pour l'intérêt qu'ils ont porté à mon travail en acceptant de rapporter mon mémoire et de participer au jury. J'ai bien conscience que ce qui nous manque le plus dans le monde académique, c'est le temps, et, quelle que soit l'intérêt de mes travaux, c'est une tâche de plus parmi les nombreuses auxquelles sont associés les professeurs. Je remercie de même Anne Fladenmuller, Stefano Secci et Sébastien Tixeuil d'avoir accepté d'être membres de ce jury.

Je suis très reconnaissant aux étudiants de Master et de doctorats successifs que j'ai encadré ou co-encadré, et dont les travaux ont permis la réalisation de ce mémoire. En particulier je pense à Thomas Silverston, maintenant Maître de conférences à Nancy ; Rabee Mustafa Abuteir, encadré avec Anne Fladenmuller et maintenant dans l'industrie ; et Kévin Vermeulen, encadré avec Timur Friedman et maintenant chargé de recherche CNRS au LAAS à Toulouse.

J'ai une attention spéciale à mes collègues du groupe de recherche Dioptra, en particulier Timur Friedman sans qui la belle dynamique de recherche dans laquelle je suis impliquée n'existerait pas, et les nouveaux thésards qui nous accompagnent : Matthieu Gouel et Berat Can Şenel. Je n'oublie pas Maxime Mouchet qui nous a rejoints en tant que postdoc et dont la contribution nous est précieuse, et tous les anciens membres qui nous ont accompagné dans cette aventure.

Un peu plus éloigné de nous, j'exprime mon immense gratitude au professeur Takumi Miyoshi du Shibaura Institute of Technology à Tokyo qui a permis une longue collaboration et toujours supporté mes travaux malgré l'éloignement.

Je remercie les membres de l'équipe NPA, du LIP6 et de l'UFR d'Ingénierie qui m'ont encouragé, et en particulier les équipes techniques dont Konstantin Kabasanov et Pierre-Emmanuel Leroux pour leur support indéfectible malgré nos expérimentations impactant parfois le réseau du labo.

Et puis il y a mon voisin de bureau, Sébastien Baey, probablement la seule personne capable de partager cet espace restreint avec moi, de supporter mon invasion de matériels informatiques et la sélection conjointe de milliers de candidatures pour le parcours RES du master informatique au cours des années.

Merci aussi aux deux directeurs du Master Informatique successifs dont j'ai été l'adjoint, Evripidis Bampis et Carlos Agon, qui ont toujours essayé de me motiver et y ont finalement réussi ! Par contre je tiens à m'excuser auprès de Valérie Chantin, notre formidable responsable administrative du master informatique, et des responsables de parcours du master informatique, ils ont eu un directeur des études un peu moins disponible lors de cette rentrée 2021.

Pour terminer, je remercie Christine, Gabriel et Anna pour leur soutien sans faille, Marie pour sa relecture minutieuse et ses conseils typographiques. Et je suis désolé de l'image négative de l'enseignant-chercheur que j'ai pu véhiculer en passant mes vacances d'été sur cette ultime épreuve de notre système académique.





# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Distribution vidéo pair-à-pair (P2P-TV)</b>	<b>5</b>
1.1 Première génération d'applications de P2P-TV . . . . .	5
1.2 Analyse des applications P2P-TV commerciales . . . . .	6
1.3 Applications P2P-TV avec mesures en plusieurs localisations . . . . .	8
1.4 Impact du comportement de l'utilisateur sur les systèmes P2P-TV . . . . .	10
1.5 Amélioration du trafic P2P-TV . . . . .	14
1.6 Conclusion du chapitre . . . . .	16
<b>2 Distribution vidéo adaptative via HTTP</b>	<b>17</b>
2.1 Dynamic Adaptive Video Streaming over HTTP (MPEG-DASH) . . . . .	18
2.2 Distribution vidéo assistée par le réseau : NAVS . . . . .	19
2.3 Distribution vidéo assistée pour le sans-fil : WNAVS . . . . .	20
2.4 Adaptation aux réseaux sans-fil d'entreprise : SMILE . . . . .	21
2.5 Amélioration pour les réseaux mobiles : DAVS . . . . .	22
2.6 Conclusion du chapitre . . . . .	23
<b>3 Cartographie de l'Internet</b>	<b>25</b>
3.1 Mesure de la topologie de l'Internet : Traceroute . . . . .	26
3.2 Amélioration de MDA Paris Traceroute : MDA-Lite . . . . .	29
3.3 Mesure de l'Internet à haute fréquence : D-Miner . . . . .	32
3.4 Cartographie de niveau routeurs : MMLTP et Limited Ltd . . . . .	35
3.5 Conclusion du chapitre . . . . .	37
<b>Perspectives</b>	<b>39</b>
<b>Bibliographie</b>	<b>43</b>



# Introduction

Si la majeure partie du trafic Internet est aujourd'hui liée à la distribution<sup>1</sup> de contenu vidéo (*video streaming*), situation confirmée pendant la crise sanitaire [San20], cela n'a pas toujours été le cas. A sa création, l'Internet n'a pas été conçu pour servir les applications multimédias, ces applications qui échangent des trafics audio/vidéo.

L'Internet est une interconnexion de réseaux échangeant des datagrammes grâce à IP (*Internet Protocol*). IP fait généralement référence à IPv4 [Pos81b], c'est-à-dire avec le numéro de version 4 pour le distinguer de sa mise à jour avec le numéro de version 6, IPv6 [DH17]. Un protocole complémentaire avec le numéro de version 5 a existé<sup>2</sup> : l'*Internet Stream Protocol* qui fut expérimenté dès 1979 [For79] pour les trafics avec contraintes temporelles (voix). Plus tard, ses évolutions ST2 et ST2+ [DB95] envisageront explicitement le transport de la vidéo, mais resteront également expérimentales. Ces protocoles sont annonciateurs de technologies plus complexes, telles qu'ATM (*Asynchronous Transfer Mode*), qui s'essaieront ultérieurement à la convergence des types de trafic [Vet95].

Dans les années 90, l'Internet devient grand public, mais l'industrie de la distribution de la vidéo commence à peine sa migration vers le numérique. Les capacités en termes de débit des réseaux informatiques sont loin d'être suffisantes pour les flux multimédias, avant même d'aborder les limitations liées aux protocoles et mécanismes utilisés. Les spectateurs accèdent aux contenus audio/vidéo via des réseaux de diffusion dédiés encore analogiques (télévision via réseaux hertziens, câble ou satellites) ou sur des médias physiques enfin numériques (le DVD, *Digital Versatile Disc*, arrive en 1995). Au même moment, de nombreux progrès autour du traitement des contenus multimédias, en particulier de la vidéo, se mettent en place afin de réduire les quantités de données à transmettre avec des codages et compressions avancés, par exemple avec les normes MPEG2/H263.

## Distribution vidéo sur le MBONE

A la fin des années 90, des technologies de réseaux avec des débits élevés font leur apparition en prenant en compte les différents types de trafic, tels que ATM. Dans le cadre d'Internet, ces services avancés resteront d'usage limité. Par contre, les besoins des utilisateurs d'Internet, principalement académiques, encouragent le développement de technologies adaptées à la distribution multimédia sur Internet. Les premières solutions sont proposées avec prise en compte des contraintes temporelles (RTP [SCFJ03]), des mécanismes de communication d'une machine vers plusieurs (IP multicast [Dee89]), et de la réservation de ressource (RSVP/Intserv [Wro97]).

C'est dans ce contexte que s'inscrivent mes premiers travaux de recherche lors de ma thèse de doctorat intitulée : « Les communications multipoints dans les réseaux haut débit multimédia : Le multicast en environnement IP sur ATM » [Fou98].

Les premiers services de distribution vidéo exploités à large échelle font leur apparition dans le monde académique, principalement avec la visioconférence, par exemple IVS de l'INRIA [Tur95]. L'usage est restreint aux utilisateurs du réseau expérimental MBONE [MB94], réseau qui connecte

---

1. Dans ce document, nous utilisons le terme « distribution » comme traduction du terme anglais « *streaming* » et non « diffusion », celui-ci nous semble trop proche du terme anglais « *broadcast* » qui a d'autres significations tant dans le domaine des réseaux que dans celui de la télévision.

2. Le protocole avec le numéro de version 5 n'est pas à proprement parler une version, mais plutôt une alternative destinée à être utilisée conjointement avec IPv4.

par des tunnels des ilots où l'extension IP multicast est disponible. La difficulté à mettre en place des communications d'un émetteur vers plusieurs destinataires à cause du développement limité d'IP multicast va amener à gérer cette fonctionnalité au niveau applicatif. C'est aussi une des motivations pour la distribution vidéo en P2P (pair-à-pair)[DLL<sup>+</sup>00, CRSZ02].

## Distribution vidéo P2P

Dans les années 2000, avec l'augmentation des débits des accès résidentiels, apparaissent de nouveaux usages d'Internet tels que la vidéo à la demande (VoD) et la télévision en direct (IPTV). Les trafics de ces nouveaux usages vers de nombreux spectateurs induisent des coûts énormes pour les diffuseurs qui doivent utiliser des infrastructures dédiées (CDN, *Content Delivery Networks*) pour optimiser leur distribution.

Ce modèle de distribution exclut les petits diffuseurs ou la diffusion personnelle. Des approches P2P sont envisagées dans ces cas, les applications P2P ayant déjà fait leurs preuves pour la distribution de fichiers. L'utilisation de ces approches ne va pas sans difficulté, car les trafics audio/vidéo imposent des contraintes temporelles absentes précédemment. Les premiers travaux présentés dans ce document concernent ainsi la P2P-TV (diffusion vidéo en direct en P2P). Nos mesures des applications P2P-TV commerciales seront parmi les premières à caractériser et comparer ces nouveaux trafics.

La distribution de contenu était une thématique du programme européen IST (*Information Society Technologies*) du FP6 (*Sixth European Framework Programme*). De 2006 à 2009, j'ai été responsable de la participation de l'UPMC au sein du réseau d'excellence IST-CONTENT du FP6. Ce projet avait pour but de fédérer une dizaine d'universités européennes autour de la distribution de contenu audio/vidéo aux utilisateurs finaux. J'ai encadré, grâce à ce projet, le doctorat de Thomas Silverston intitulé « *Peer-to-peer Video Live Streaming : Measurement and Traffic Analysis* » [Sil09].

Nous avons contribué à la mesure et à l'analyse des approches P2P pour la distribution vidéo que nous avons développées antérieurement et enrichies avec nos partenaires européens. Puis, dans le cadre de la collaboration franco-japonaise CNRS/WIDE, nous avons développé des mesures à large échelle de ces applications émergentes. Les résultats obtenus nous ont conduits à proposer de nouveaux mécanismes spécifiques au P2P-TV, en particulier au niveau de l'incitation des utilisateurs à participer afin de maintenir des ressources dans le réseau P2P virtuel utilisé.

Les applications P2P-TV commerciales sont principalement originaires de Chine et fortement utilisées en Asie. Nous avons donc prolongé nos collaborations au Japon, en particulier avec le professeur Takumi Miyoshi au Shibaura Institute of Technology. Nos travaux ont abordé l'impact du comportement des utilisateurs et les possibilités de contrôle du trafic généré par ces systèmes P2P. La caractérisation du comportement des utilisateurs est fondamentale dans le cadre des réseaux P2P puisque l'utilisateur n'est pas seulement consommateur du contenu, mais aussi acteur du système de distribution.

## Distribution vidéo via HTTP

Lors de la dernière décennie, tous les services impliquant de la vidéo (télévision traditionnelle, vidéo à la demande, diffusion pour mobiles...) ont commencé à converger vers Internet. Ces contenus multimédias fournis via différents réseaux sont destinés à des spectateurs accédant à la ressource avec des équipements multiples (ordinateur, télévisions, *smartphones*, tablettes, consoles de jeux...). Cette hétérogénéité massive en termes de réseaux d'accès, de terminaux et d'attentes des utilisateurs nécessite des solutions efficaces, interopérables et universelles. C'est à travers le protocole HTTP (Hypertext Transfer Protocol) que cette convergence a été réalisée, et ce malgré les réticences longtemps éprouvées face à l'utilisation de ce protocole pour les

flux audio/vidéo. En effet, HTTP repose sur la couche de transport TCP qui a longtemps été perçue comme inadaptée pour les trafics avec contrainte temporelle à cause de ses mécanismes de retransmission et de contrôle de congestion.

L'utilisation de HTTP à l'avantage de réutiliser une infrastructure de distribution efficace, avec un protocole qui traverse les filtres (*firewall*) et les modifications d'adresses (*NAT, Network Address Translation*). Avec le développement d'applications adaptatives qui gèrent de manière efficace les tampons mémoires et s'adaptent à la forme du trafic résultant de TCP, les trafics vidéo sont aujourd'hui en grande partie transmis via HTTP.

Le changement vers HTTP a été principalement impulsé par les solutions propriétaires de distribution vidéo. Si l'on considère une première génération avec des lecteurs dédiés, HTTP était utilisé pour accéder aux contenus de manière progressive, mais non adaptative. La génération actuelle utilise l'adaptation aux conditions du réseau, ou HAS (*HTTP Adaptive Streaming*), introduit dès 2009 par Apple avec sa proposition HLS (*HTTP Live Streaming*) [PM17]. Les autres acteurs du domaine ont ensuite fait d'autres propositions jusqu'à ce que MPEG propose DASH (*Dynamic Adaptive Streaming over HTTP*) en 2011 [AI11].

Ces solutions adaptatives reposent sur des mécanismes de sélections exécutés au niveau du lecteur du client afin de choisir sur le serveur HTTP les données adaptées parmi plusieurs qualités d'un même contenu. La problématique de recherche se focalise sur cette adaptation qui introduit un niveau de dynamique supplémentaire dans la chaîne de distribution et qui joue sur la perception de la QoE (Qualité d'expérience) du spectateur [TLPR14]. Les choix effectués sont cruciaux vu l'importance de la part du trafic vidéo (70% du trafic Internet en 2017 d'après Cisco [Cisb]) et l'adoption massive du HAS par les fournisseurs OTT (*Over-the-Top*), c'est-à-dire les fournisseurs de contenu en ligne sur Internet (Netflix, YouTube, Amazon Prime...) [Bin15].

En 2014, en combinant mon expérience de la distribution de contenu audio/vidéo et celle des réseaux d'accès sans-fil d'Anne Fladenmuller, nous avons co-encadré le doctorat de Rabee Abuteir intitulé « *Managing Adaptive Video Streaming in Edge Networks* ». Ce travail propose une amélioration de la distribution des flux DASH à l'aide des technologies SDN/NFV, qui prend tout son sens vu le développement de ces technologies tant au niveau de la bordure du réseau pour contrôler le trafic proche du client que d'une manière plus générale avec la virtualisation de la chaîne de traitement des contenus multimédias.

## Vers la cartographie de l'Internet

Mes travaux sur la P2P-TV se terminaient sur des solutions visant à mieux localiser les pairs afin de contrôler la proximité des trafics. La distribution de vidéo via HTTP pose le problème de la localisation des serveurs utilisés pour optimiser leur choix. Ces deux approches de distribution vidéo bénéficieraient d'obtenir des informations de localisation pour optimiser leurs services.

En 2016, j'ai eu l'opportunité de travailler avec Timur Friedman sur les mesures de l'Internet. Nous avons démarré ensemble le projet « Cartographie de l'Internet » financé par le ministère des Armées. Celui-ci nous a permis de financer, entre autres, le travail de doctorat de Kevin Vermeulen dont la thèse est intitulée « *Improved algorithms for capturing Internet maps* » [Ver20].

La réalisation des cartes à l'échelle de l'Internet repose sur l'assemblage de nombreuses mesures de routes entre deux points du réseau. Historiquement, l'outil de base est Traceroute [Jac88] pour des informations au niveau des adresses IP des interfaces rencontrées sur une route. Mais cet outil ne tient pas compte de l'évolution de l'Internet qui comporte aujourd'hui de nombreuses routes multi-chemins [Hop00, AFT11, VSFF18], où le trafic vers une même destination est réparti sur plusieurs liens (*load balancing*). Des travaux antérieurs ont développé des techniques de mesure pour découvrir ces routes [AFT07], mais qui n'ont été ni déployées en production ni utilisées à l'échelle d'Internet, en grande partie en raison du nombre de sondes et de la durée nécessaires. En conséquence, les cartes d'Internet disponibles [CAI19, rip] sont incomplètes : nos mesures

découvrent 64% de liens supplémentaires.

Construire des cartes précises et complètes est crucial pour étudier la résilience aux pannes et aux attaques de l'Internet [Sch<sup>+</sup>11, LB17], mais aussi pour supporter les applications comme la distribution de contenu ou la géolocalisation [KBJK<sup>+</sup>06] par exemple.

En collaboration avec Robert Beverly et Justin Rohrer de NPS (*Naval Postgraduate School*), nous avons développé « *Diamond Miner* », l'outil de recensement d'Internet à large échelle le plus performant actuellement à partir d'un seul point de mesures. Il envoie rapidement des sondes de manière aléatoire, en ordre permuté pour éviter de surcharger une seule route, et complète itérativement sa vision du réseau [VRB<sup>+</sup>20]. Cet outil est le premier à offrir une visibilité sur les répartiteurs de charge de plus en plus présents dans l'Internet actuel.

Pour compléter ces travaux, et obtenir des cartes au niveau routeur et non des adresses IP, nous avons travaillé à la résolution d'alias qui essaie de regrouper des interfaces. Nous avons également collaboré avec le professeur Reza Rejaie de l'Université d'Oregon sur cette technique et proposé une nouvelle approche basée sur la limitation en débit des sondes ICMP [VLA<sup>+</sup>20].

## Plan et contributions

Suite à la présente introduction, ce mémoire se compose de trois chapitres de synthèse scientifique qui décrivent l'évolution de mon travail de recherche, et se termine par des perspectives de recherche.

Le **chapitre 1** concerne la mesure et l'analyse des applications P2P-TV pour la distribution vidéo. Les contributions présentées sont relatives aux applications propriétaires les plus utilisées à la fin des années 2000. Avec principalement Thomas Silverston, nous avons proposé les premières études comparatives de ces applications [SF07, SFS<sup>+</sup>09]. Nos collaborations au Japon nous ont permis de proposer des comparaisons entre la France et l'Asie [SFSC07, SFSC10], d'évaluer l'équité et la localité des échanges [SJCA<sup>+</sup>11], puis de proposer un mécanisme d'incitation adapté au P2P-TV [SFC08]. Les collaborations avec le professeur Takumi Miyoshi nous ont permis de proposer des analyses du comportement des utilisateurs [NMF11] puis des analyses d'impact des modifications du trafic vers certains pairs : ajout de délais [SMF12], limitation du débit [HVMF14], et introduction de pertes [HVMMF13]. Une dernière contribution propose l'analyse du comportement des utilisateurs en mesure passive sur plusieurs canaux [GFTM16].

Le **chapitre 2** se penche sur le modèle de distribution dominant actuel basé sur la distribution de la vidéo HAS. Les contributions présentées ont été réalisées dans le cadre d'une collaboration avec Anne Fladenmuller pour l'encadrement de la thèse de Rabee Mustapha Abuteir. Nous avons tout d'abord proposé une approche basée sur un contrôle centralisé via SDN pour améliorer la distribution vidéo dans les réseaux résidentiels filaires [AFF16b], nous avons développé son extension pour les réseaux sans-fil résidentiels [AFF16a] et pour les réseaux d'entreprise [AFF18]. Avec le professeur Mostafa Ammar, nous avons réalisé une dernière contribution sur l'adaptation du débit dans le cadre des réseaux sans-fil [AFFA17].

Le **chapitre 3** aborde des problématiques liées à la cartographie de l'Internet. Les contributions présentées ont toutes été réalisées principalement avec Timur Friedman et Kévin Vermeulen. Une première réalisation a consisté à optimiser l'outil MDA Paris Traceroute pour fournir les routes les plus précises et complètes possible [VSFF18]. Pour passer à l'échelle de l'Internet, nous avons collaboré avec Robert Beverly et Justin Rohrer et proposé l'outil Diamond Miner qui est à ce jour le plus complet et le plus rapide pour cartographier l'Internet [VRB<sup>+</sup>20]. Pour compléter ces travaux, nous avons également collaboré avec le professeur Reza Rejaie sur une nouvelle approche de la résolution d'alias basée sur la limitation en débit des sondes ICMP [VLA<sup>+</sup>20].

Enfin, la dernière partie de ce mémoire présente les perspectives de ces travaux de recherche pour les prochaines années.

# 1 Distribution vidéo pair-à-pair (P2P-TV)

Au milieu des années 2000, la généralisation des accès résidentiels à haut débit démultiplie le nombre d'utilisateurs du réseau Internet et la demande de contenus multimédia. Les capacités de transmission nécessaires requièrent des infrastructures coûteuses qui peinent à satisfaire la demande.

Au même moment, de nouveaux usages apparaissent avec des utilisateurs qui, de simples consommateurs avec des navigateurs passifs, deviennent des producteurs de contenu : ils mettent à disposition de nouveaux contenus et forment des communautés dans lesquelles la participation est active, et où l'interaction et le partage d'informations sont fortement encouragés.

Les demandes des utilisateurs et les nouvelles formes d'interaction entraînent l'évolution des services attendus, en particulier des capacités de distribution de contenus dynamiques et décentralisées. Le paradigme P2P (pair-à-pair), déjà massivement utilisé pour échanger des fichiers [LKR06, GDS<sup>+</sup>03], semble à ce moment très prometteur pour des contenus multimédia avec contraintes temporelles tels que les trafics vidéo, et plus spécifiquement ceux de la télévision en direct ou P2P-TV [SMG<sup>+</sup>07].

Les travaux présentés dans la suite ont eu pour principal soutien le projet européen FP6 IST NoE CONTENT de 2006 à 2009 dont j'étais responsable pour l'UPMC. Le but de ce réseau d'excellence était de fédérer autour de la thématique « réseaux de distribution de contenu pour les utilisateurs résidentiels » les partenaires européens suivants : Université Carlos III de Madrid (coordinateur), Université Politechnique de Catalogne, Université de Lancaster, Université de Coimbra, NKUA d'Athènes, Université Technique de Darmstadt, AGH de Cracovie, Université d'Oslo, Université Technique de Delft et l'Université de Naples. Ce projet a permis de financer la thèse de Thomas Silverston que j'ai encadrée de 2005 à 2009 sur la thématique du P2P-TV et avec les contributions suivantes [SF06, SF07, SFS07, SFS08, SFC08, SFS<sup>+</sup>09].

Nos travaux ont aussi bénéficié de nos collaborations au Japon pour les mesures multipoints avec Kenjiro Cho du IJ Research Lab dans le cadre du projet WIDE-CNRS [SFSC07, SFSC09, SFSC10, SJCA<sup>+</sup>11] et de la coopération avec le professeur Takumi Miyoshi du Shibaura Institute of Technology qui a obtenu une bourse JSPS KAKENHI sur notre thématique pour financer des étudiants co-encadrés au Japon : Yuko Nakamura [NMF11], Yuki Shinozaki [SMF12], Hiep Hoang-Van [HVMF13, HVSMF14, HVMF14] et Koki Mizutani [HVMMF13, MMF15].

Ces travaux ont été complétés par des projets internes des laboratoires LIP6 (SU) et LTCI (Télécom ParisTech) finançant les stagiaires de M2 suivants : Aikaterini Andreadou, Jingxian Lu, Duyen Hoa Ha, Thanh Dat Phung, Wang Man Xue, Trinh Thanh Su, Pierre Vigreux, Valentin del Piccolo, Naoki Takeuchi et Marwan Ghanem. Deux de ces stages ont conduit à des publications [WNMF12, GFTM16].

## 1.1 Première génération d'applications de P2P-TV

Au début des années 2000, de nombreux travaux académiques sont initiés pour étendre ceux déjà réalisés sur le P2P ou simplement pour pallier l'absence de *multicast* globalement disponible au sein de l'Internet. Plusieurs approches ont été proposées que l'on peut regrouper en deux types de fonctionnement P2P : construction d'une surcouche (*overlay*) soit « orienté arbre », soit « orienté maillage ».



Les approches « orientées arbre » (*tree-based* ou *source driven*) reposent sur un graphe de distribution arborescent ayant pour racine la source du contenu d'où sont poussées les données. Le fonctionnement d'un pair, ou nœud, repose sur la réception des données de nœuds parents, qui peuvent être la source ou d'autres pairs intermédiaires. Si la structure est relativement stable (si les pairs ne changent pas trop de configuration), ces approches sont assez efficaces, car les données sont poussées de pair en pair sans nécessiter de signalisation explicite. Si la dynamique des pairs augmente (taux de désabonnement élevé par exemple), les pairs doivent temporiser leur transmission afin d'adapter l'arborescence assez rapidement pour éviter la perte de données. On peut citer les travaux Narada [CRSZ02], NICE [BBK02] ou encore Splitstream [CDK<sup>+</sup>04].

Les approches « orientées maillage » (*mesh-based* ou *data-driven*) reposent sur la gestion par chaque pair d'un sous-ensemble de pairs pour obtenir un certain bloc de données. Chaque pair doit savoir quels morceaux appartiennent à quels pairs et demander explicitement les blocs dont il a besoin. Ce type d'approche implique une signalisation importante, due d'une part à l'échange de table de contenu des tampons de données entre les pairs et d'autre part au processus de demande explicite de blocs de données. Ces approches « orientées maillage » offrent une bonne résilience aux défaillances, mais nécessitent des tampons mémoires importants pour récupérer les blocs de données. On peut citer les travaux CoolStreaming/DOnet [ZLLY05] et les nombreuses applications commerciales qui vont s'en inspirer et que nous étudions dans la suite.

### 1.2 Analyse des applications P2P-TV commerciales

Parallèlement à ces propositions académiques, plusieurs applications commerciales suscitaient déjà un fort engouement. Ces applications propriétaires indiquaient utiliser une approche « orientée maillage » comme DOnet. Dans ces systèmes P2P, les données sont divisées en blocs (*chunks*) et chaque pair échange les informations sur les blocs dont il dispose. Ensuite, chaque pair peut télécharger des blocs de données de plusieurs pairs simultanément. Cependant, le mode de fonctionnement exact de ces applications émergentes était encore largement méconnu. Dans ce contexte, il était important d'évaluer l'impact du trafic de ces applications de manière similaire aux nombreuses études menées sur les systèmes de partage de fichiers.

Nous avons profité de la coupe du monde de football 2006 pour mesurer le trafic généré pour visualiser un même contenu par plusieurs applications P2P-TV commerciales : PPLive [PPL06], PPStream [PPS06], SOPCast [SOP06] et TVAnts [TVA06]. Nous avons choisi ces applications, car elles étaient les plus populaires en 2006. Nous avons réalisé des mesures passives, les applications propriétaires étant de type boîte noire : les mécanismes internes n'étaient pas connus et nous ne pouvions que les inférer aux vues des trafics réseau observés. Ces premières expérimentations ont eu lieu pendant un événement planétaire afin de maximiser la participation à grande échelle avec un grand nombre d'utilisateurs (et donc de pairs). Grâce aux données collectées à partir du trafic échangé par un pair, nous avons caractérisé les différents modèles de trafic sur des aspects quantitatifs (nombre de pairs observés, quantité de données échangées, durée de participation...) et qualitatifs (ports, protocoles, taille des paquets, temps d'inter-arrivé...).

Nos résultats montrent qu'à partir du même point de mesure, les applications génèrent des modèles de trafic très différents. Cela peut être lié à la configuration des participants, mais probablement surtout à des mécanismes différents pour le maintien du voisinage de pair et pour distribuer la vidéo. La figure 1.1 présente une des premières courbes comparatives entre les trafics des quatre applications étudiées.

Une première analyse a consisté à classer les échanges soit en trafic de signalisation, soit en trafic de données vidéo, à l'aide d'heuristiques plus ou moins complexes. La figure 1.2 montre la possibilité de faire une discrimination en observant la densité de distribution par rapport au temps d'inter-arrivée et à la taille des paquets échangés dans une capture de trafic PPLive. Deux

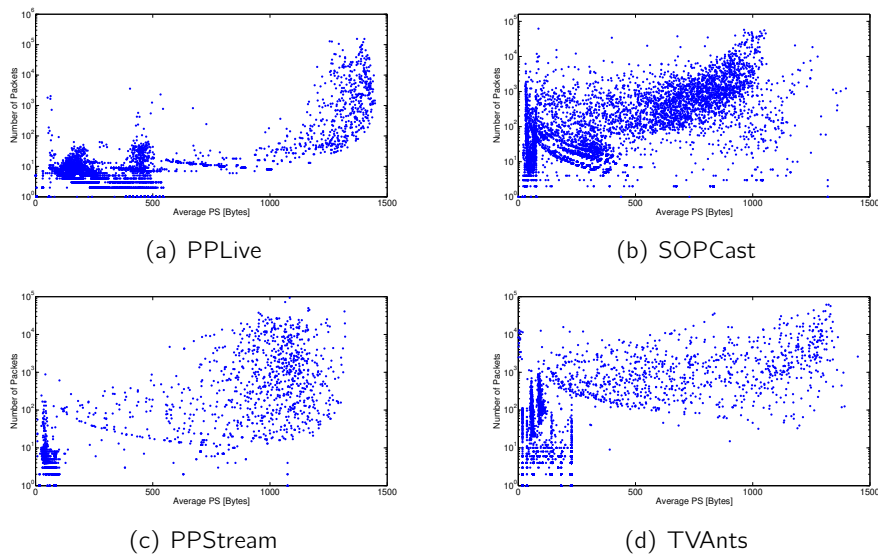


Figure 1.1 – Flux montant : nombre de paquets émis par taille moyenne de ces paquets

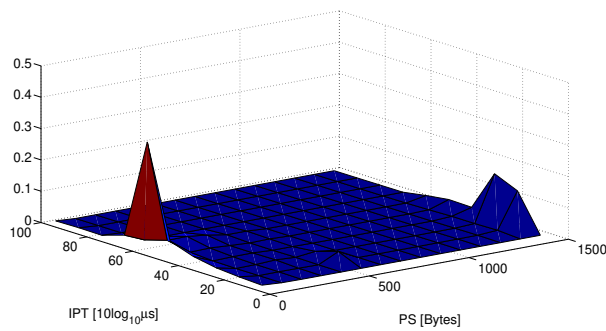


Figure 1.2 – Distribution du temps inter-paquet et de la taille de paquet pour PPLive.

pics sont observés : l'un composé de paquets de petite taille assez espacés dans le temps (la signalisation) et l'autre avec des paquets de grande taille et rapprochés dans le temps (les données vidéo). L'utilisation de cette information permet de séparer les trafics sur une trace de débit de l'application PPLive présentée dans la figure 1.3.

Nos résultats ont été présentés dans l'article « *Measuring P2P IPTV Systems* » accepté à la conférence ACM NOSSDAV 2007 [SF07]. Cette conférence, parrainée par les groupes SIGCOMM et SIGMM de l'ACM, fait référence pour la distribution vidéo sur Internet et a assuré à nos travaux une bonne visibilité avec plus de 200 citations. Cette publication a été complétée par l'article « *Traffic Analysis of Peer-to-Peer IPTV Communities* » publié dans la revue Elsevier Computer Networks [SFS<sup>+</sup>09] avec plus de 100 citations, confirmant l'antériorité et l'intérêt de ces mesures. Même s'il existait déjà des études qui montraient que la diffusion en direct à grande échelle pouvait être prise en charge par une approche P2P[SGMZ04], et que quelques propositions comme celle de DOnet [ZLLY05] avait été implémentées, notre travail était le seul qui comparait plusieurs applications et essayait d'inférer le comportement de tels systèmes P2P.

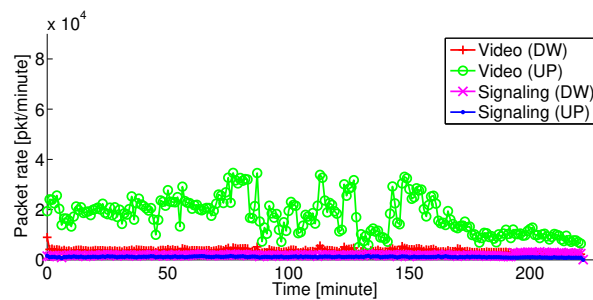


Figure 1.3 – Débit des trafics de données vidéo et de signalisation entrant et sortant pour PPlive (bin de 60s)

### 1.3 Applications P2P-TV avec mesures en plusieurs localisations

Les précédents travaux ne disposaient que d'un seul point de mesure pour étudier le trafic des applications P2P-TV utilisées mondialement. Dans le cadre de collaborations avec des partenaires au Japon, nous avons introduit la disparité géographique des pairs comme paramètre supplémentaire pour estimer l'impact de l'endroit où se situent les utilisateurs sur l'organisation des applications P2P mesurées.

Ainsi, nous avons effectué de nouvelles expériences de mesure à plus large échelle afin d'obtenir une vision étendue du réseau P2P. Nous avons mesuré le trafic d'applications P2P-TV en France et au Japon (figure 1.4) afin d'étudier l'organisation géographique du réseau P2P et les échanges de trafic entre pairs.

Les nouveaux aspects pris en compte sont : la contribution et la réciprocité des échanges nécessaires pour le bon fonctionnement des systèmes P2P ; la proximité entre les pairs permettant une meilleure efficacité et une réduction des coûts pour les FAI. Pour ces nouvelles campagnes de mesures passives, nous avons choisi les applications les plus populaires en 2007 : PPStream, TVUPlayer, SOPCast et TVAnts (nous avons remplacé PPlive par TVUPlayer [TVU11]).

Nous avons effectué notre campagne de mesures au cours de l'été 2007, car nous avons accès au réseau WIDE situé à l'université de Tokyo au Japon. Pour chaque expérience, la même application P2P-TV était lancée sur toutes les machines. Nous avons mesuré des événements sportifs et plus particulièrement des matchs de football, car ils présentent un fort intérêt à être regardé en direct. Une importante communauté d'utilisateurs les regarde via les systèmes P2P-TV.

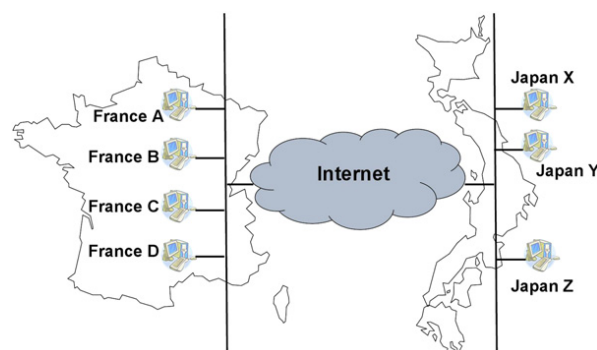
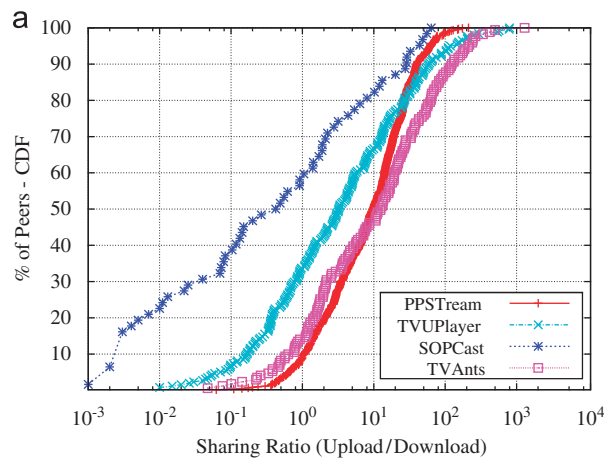


Figure 1.4 – Expérimentation à grande échelle à partir du Japon et de la France

Figure 1.5 – Proportion de trafic partagé (*upload/download*)

Une première analyse s'intéresse à la collaboration entre les pairs à travers la proportion de trafic partagé (*sharing ratio*), c'est-à-dire la quantité de trafic envoyé par un seul pair (*upload*) divisée par la quantité de trafic reçu par ce même pair (*download*). La figure 1.5 présente la fonction de distribution cumulative (CDF) du trafic partagé entre les pairs étudiés et les autres pairs avec lesquels ils communiquent. Quand la proportion de trafic partagé est supérieure à 1, les pairs étudiés sont altruistes et transmettent plus de données qu'ils n'en reçoivent. Quand cette proportion vaut 1, alors l'échange de trafic est équilibré. Et quand la proportion est inférieure à 1, alors les pairs étudiés ont davantage reçu d'autres pairs. À part pour SOPCast, les pairs des autres applications semblent plus consommateurs que producteurs. C'est probablement en partie dû au débit d'accès important dont nous disposons sur les campus universitaires. Cela pose aussi la question de l'adaptation à l'échelle des applications en cas de forte asymétrie. Si les échanges entre les pairs ne sont pas réciproques, certains pairs, ou la source elle-même, vont devoir compenser et on s'éloigne des bénéfices recherchés avec l'approche P2P. Une solution pour remédier à ce type de problème est d'intégrer des mécanismes d'incitation de manière similaire à ceux utilisés dans BitTorrent [Coh03] pour le partage de fichiers. Nous avons contribué dans ce sens avec l'article « *Towards an Incentive Mechanism for Peer-to-Peer Multimedia Live Streaming Systems* » accepté à la conférence IEEE P2P 2008 [SFC08].

De nos traces, nous avons extrait des mesures qui nous semblaient représentatives pour étudier la localisation géographique des pairs. Par exemple, dans la figure 1.6, nous avons mesuré l'application PPSTream avec cinq ordinateurs (pairs) : trois étaient localisés en France et deux au Japon. Nous avons effectué la correspondance entre les adresses IP des pairs observés et les pays avec la base de données GeoIP [Geo], puis présenté sous forme d'histogrammes empilés pour chaque direction (*upload/download*) la proportion de pays en fonction du nombre de pairs communicants avec les pairs mesurés puis de la quantité de trafic entre ces mêmes pairs (seuls les pays qui participent à plus de 1% sont listés).

L'analyse de ces résultats doit tenir compte de l'origine des applications utilisées (les sources des données provenaient de serveurs en Chine) et des utilisateurs fortement localisés en Asie avec ici un contenu motivant plutôt pour des utilisateurs de France et de Grande-Bretagne (match Liverpool-Toulouse). On observe effectivement que parmi les pairs impliqués, plus de la moitié se situent en Asie (surtout Chine et Hong-kong) : la répartition est relativement similaire en *upload* et en *download* et doit correspondre principalement à la signalisation entre pairs (peu de différence entre les pairs contrôlés en France et au Japon). La quantité de données échangées va

## 1 Distribution vidéo pair-à-pair (P2P-TV)

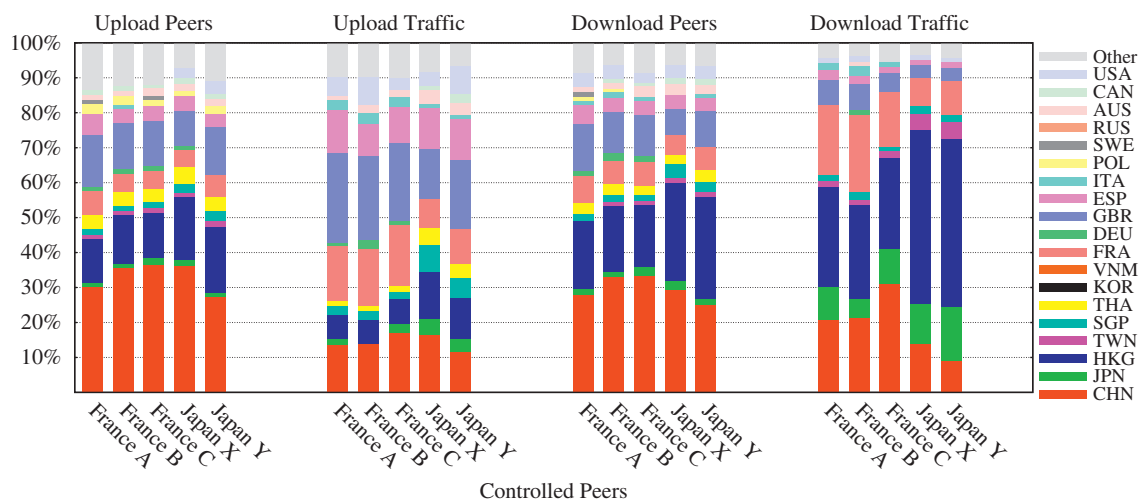


Figure 1.6 – Localisation géographique des pairs et du trafic pour PPStream

refléter la transmission du flux vidéo et va provenir plutôt des pairs proches de la source asiatique (ici on observe surtout Hong-kong et le Japon, probablement grâce aux accès à haut débit). Pour la redistribution, les trafics vont plus vers des pairs en France, Grande-Bretagne, Espagne et USA (à noter un renforcement vers Singapour et la Thaïlande pour les pairs mesurés au Japon).

Une étude comparative des différentes applications P2P-TV sur plusieurs événements sportifs confirme nos précédentes observations d'asymétrie du trafic vidéo. Nous avons également étudié les distances d'accès aux autres pairs (nombre de sauts ou d'AS) et constaté que les pairs français téléchargent le trafic plus loin que les pairs japonais et le ré-émettent plus proches d'eux. Les mécanismes utilisés ne semblent pas insister sur la réciprocité, mais plutôt sur une approche opportuniste favorisant les accès avec les débits les plus élevés.

Nos travaux de mesure en France et au Japon ont été initialement présentés au student workshop de la conférence ACM CoNEXT 2007 [SFSC07]. Les analyses de l'équité et de la distribution géographique des pairs ont été présentées dans l'article « *On Fairness and Locality in P2P-TV through Large-Scale Measurement Experiment* » à la conférence IEEE Globecom 2010 [SFSC10], puis publiées dans la revue Elsevier Computer Networks [SJCA<sup>+</sup>11].

### 1.4 Impact du comportement de l'utilisateur sur les systèmes P2P-TV

Au début des années 2010, les systèmes P2P-TV suscitent toujours un fort intérêt. Les études menées sur ces applications propriétaires concernent principalement l'organisation des pairs et les caractéristiques des trafics échangés puisque les mécanismes internes sont inaccessibles.

Dans son article de 2006 [HLL<sup>+</sup>06], Hei constatait que les utilisateurs de P2P-TV avaient des comportements proches de ceux de la télévision classique. Une étude détaillée d'un système IPTV dans [CRC<sup>+</sup>08] caractérisait plus en détail le comportement des utilisateurs avec en particulier la notion de changement de canal (les systèmes de télévision proposent généralement de nombreux canaux ou chaînes). Les systèmes P2P-TV que nous avons étudiés proposaient également de nombreux canaux, chaque canal correspondant probablement à un overlay intégrant les pairs souhaitant recevoir son contenu. Les pairs peuvent changer de canal à tout moment, ce qui ajoute un facteur dynamique supplémentaire. C'est cet aspect que nous avons l'intention d'aborder en

contrôlant les canaux choisis lors des mesures.

En 2011, nous avons à nouveau sélectionné les applications P2P-TV les plus populaires : SOP-Cast, PPSTream, PPLive, et UUsee (nous avons remplacé TVAnts et TVUPlayer par PPLive déjà présenté précédemment et introduit UUsee [UUS07]) pour de nouvelles campagnes de mesures passives. Ces quatre applications originaires de Chine diffusaient tous les canaux de télévision nationale en direct. Nous avons donc réglé chaque application sur les cinq canaux les plus suivis : CCTV1, CCTV2, CCTV4, CCTV10 et CCTV13.

Nous avons défini plusieurs comportements d'utilisateurs, introduisant le changement de canal, voire les changements successifs (*zapping*), afin d'observer la réaction des applications lors de ces événements transitoires. Dans la suite, nous présentons un scénario où un utilisateur SOAPCast se connecte au système P2P-TV, puis après 30 secondes choisit successivement cinq canaux et reste sur le dernier. Une attente de 120 secondes est réalisée pour atteindre un état considéré stable servant de référence pour comparer la situation dans les phases transitoires.

La figure 1.7 présente le chronogramme décrivant l'expérience. Le temps d'attente sur chaque canal  $t_s$  doit être assez rapide pour satisfaire l'utilisateur, mais suffisamment long pour que l'application puisse contacter les pairs, récupérer assez de données pour remplir son tampon de lecture et décoder le début du contenu vidéo. La valeur présentée ici est  $t_s = 25s$ , car l'apparition de la vidéo prenait en moyenne de 15 à 20 secondes. Pour la dernière chaîne, le temps d'écoute est également de 200 secondes.

Pour illustrer la surcharge en phase transitoire, la figure 1.8 présente les pairs contactés. Le code couleur indique le moment où les pairs sont apparus. En rouge la totalité, puis les autres couleurs correspondent aux pairs impliqués dans les échanges associés aux différents canaux. La figure 1.9 procède de même, mais avec les trafics correspondants.

Ces travaux sur le comportement des utilisateurs en phase transitoire sont importants pour évaluer l'impact probable des systèmes P2P-TV sur le réseau. En l'absence de mécanismes de régulation dédiés, si la demande de ressource ponctuelle engorge les réseaux d'accès et que ces applications répondent en contactant davantage de pairs voisins, le phénomène de congestion sera démultiplié. Si l'on ajoute à cela des utilisateurs qui s'impatientent et changent encore plus de canaux, la situation risque de se dégrader encore plus.

Ces travaux ont été publiés au congrès WTC 2012 [WNMF12] et les nombreuses questions auxquelles nous avons eu à répondre furent sur la validation du comportement des utilisateurs et l'observation de pairs présents sur plusieurs canaux, information que nous n'avions pas. En 2013, nous avons eu l'opportunité de travailler en ce sens avec une nouvelle campagne de mesures pour l'analyse multicanale. L'utilisation de sondes (pairs abonnés de manière permanente à un canal) permet l'observation de l'apparition et le suivi de pairs sur un ou plusieurs canaux. Évidemment la vision est partielle, mais cela permet d'avoir une idée de la dynamique et des types de pairs présents en analysant le trafic des sondes.

Nous nous sommes focalisés sur une seule application P2P-TV : PPTV (ex-PPLive), mais avec plus de canaux (les neuf plus populaires de ce système). Notre but était de distinguer des pairs de contrôle qui présentent surtout un trafic de signalisation, et des super-pairs qui redistribuent sur plusieurs canaux (éventuellement des serveurs vidéo contrôlés par les propriétaires de l'application, les services commerciaux de distribution vidéo P2P ayant tendance à devenir hybride pour garantir un minimum de qualité de service).

Sur des captures longues (12 heures), on observe la présence de 5 à 12% de pairs multicanaux par canal sur le nombre total de pairs. Par exemple, la figure 1.10 présente le trafic sur l'un des canaux en fonction des pairs qui apparaissent successivement dans la capture. Les lignes orange montrent le trafic de pairs à canal unique (n'apparaissant que sur ce canal), et les lignes bleues montrent celui de pairs multicanaux (apparaissant aussi sur d'autres canaux).

Deux types de comportements sont identifiés pour les pairs multicanaux : la figure 1.11 présente

# 1 Distribution vidéo pair-à-pair (P2P-TV)

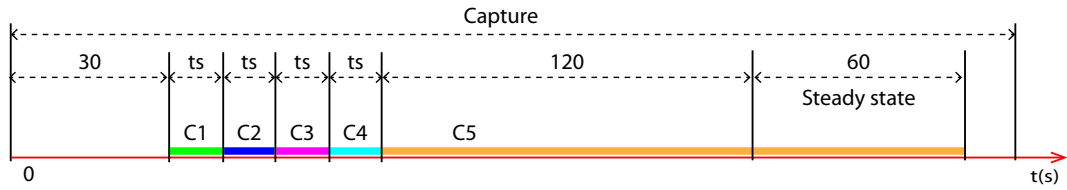


Figure 1.7 – Scénario avec 5 canaux P2P-TV

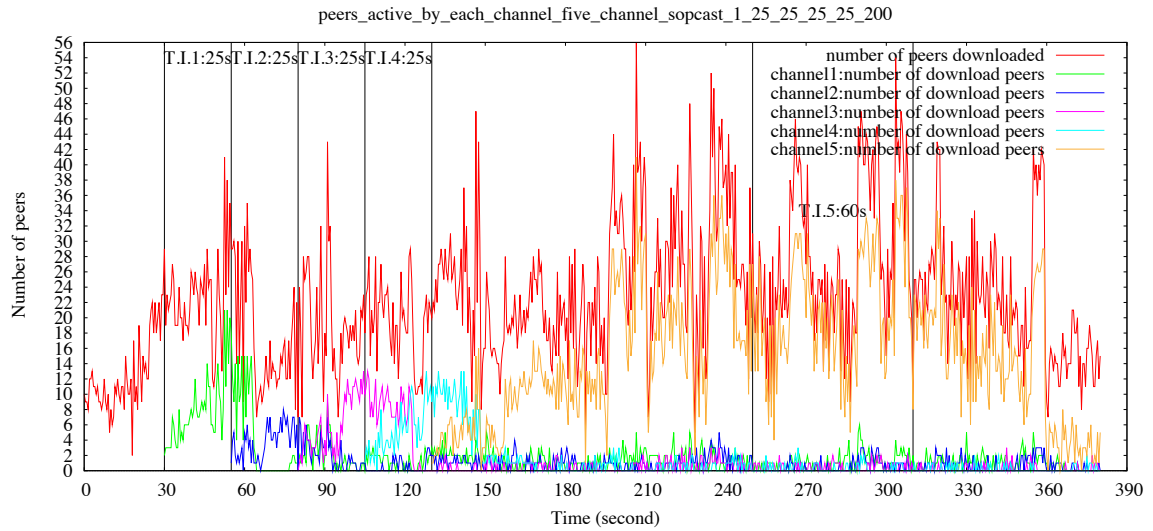


Figure 1.8 – Pairs actifs par canaux avec SOAPCast lors d'un zapping sur 5 canaux.

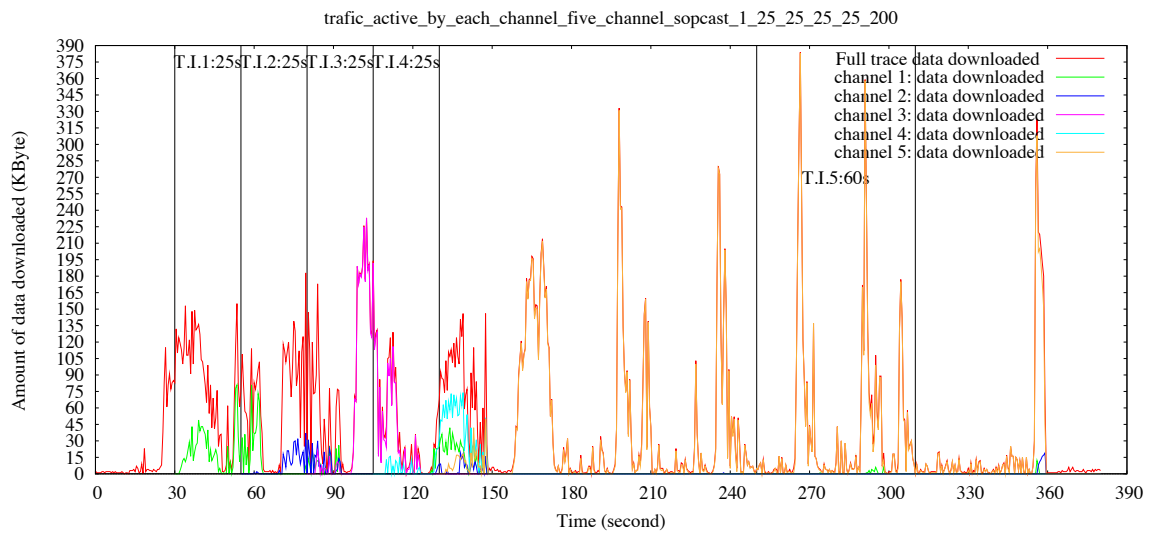


Figure 1.9 – Trafic par canaux avec SOAPCast lors d'un zapping sur 5 canaux.

## 1.4 Impact du comportement de l'utilisateur sur les systèmes P2P-TV

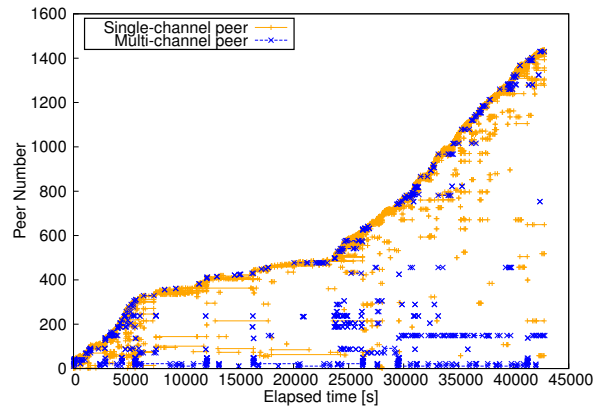


Figure 1.10 – Tous les flux sur un canal PPTV.

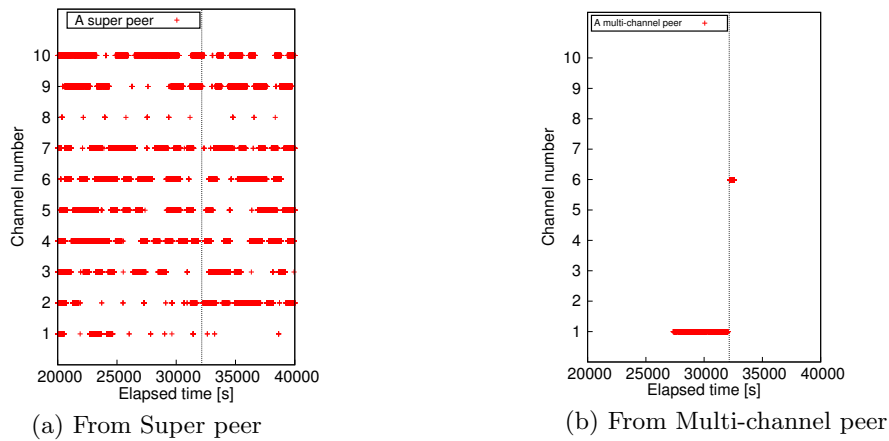


Figure 1.11 – Flux d'un pair apparaissant sur plusieurs pairs de mesure.

les trafics sur chaque canal de ces pairs. Sur (a), les sondes ont observé simultanément le même pair parmi plusieurs canaux pendant une longue durée. Nous supposons que c'est un super-pair, qui fournit des services de distribution vidéo stables en se consacrant à l'envoi de paquets vidéo. Au contraire, sur (b), on observe une simple transition séquentielle d'un canal à un autre. Ce pair multicanal n'a pas été observé simultanément par les sondes et a rapidement disparu. Nous supposons que ce pair est celui d'un utilisateur qui est associé uniquement à un unique canal à la fois.

Nos travaux ayant suscité l'intérêt de collègues en analyse de graphes de terrain, nous avons relancé en 2015 une campagne de mesures pour approfondir le suivi de la présence de pairs actifs sur différents canaux. Ce deuxième jeu de données permet de comparer l'évolution des situations. Ainsi les pairs présents sur plus d'un canal passent de 8% à 16% du nombre total de pairs observés. Ce nombre est seulement de 2% et 3% si l'on se limite aux pairs actifs (impliqués dans le trafic de contenu vidéo). De plus, l'introduction d'analyses basées sur des fenêtres temporelles glissantes a permis de suivre précisément les pairs passant d'un canal à l'autre et de mieux identifier les super-pairs.

Ces résultats sont intéressants pour caractériser le comportement des pairs et extrapoler le



comportement des utilisateurs. Ils peuvent être utiles pour les fournisseurs de contenu d'adapter l'offre vidéo et pour les opérateurs de réseaux d'optimiser leur infrastructure (problématique abordée dans la prochaine section).

Ces analyses multicanales passives ont été publiées initialement en 2015 [MMF15] et les travaux complémentaires en collaboration avec Fabien Tarissan ont obtenu un *best paper award* à la conférence IEICE/IEEE APNOMS 2016 [GFTM16].

### 1.5 Amélioration du trafic P2P-TV

Depuis l'apparition des premières applications de partage de fichiers P2P, le trafic lié aux applications basées sur l'approche P2P a posé problème aux opérateurs [KRP05]. Outre sa quantité, sa nature imprévisible et ses chemins non optimisés posaient régulièrement la question de sa régulation au dépens de la neutralité du réseau. Avec la démultiplication des trafics liés aux contenus multimédia, le contrôle de ces trafics se posait à nouveau au début des années 2010.

Les communications P2P sont routées au niveau des pairs, gérées dans la surcouche applicative (*overlay*) construite en s'abstrayant de la topologie réelle des infrastructures. Il en résulte des échanges de données non optimaux avec potentiellement de nombreux trafics coûteux et non nécessaires entre les AS et les différents opérateurs [PCB06]. En réponse, les opérateurs pourraient essayer de limiter ces trafics, voir les bloquer, solution temporaire avant que ces applications n'évoluent et ne masquent leur trafic, rendant leur régulation encore plus complexe.

Gérer ce problème de localisation du trafic revient à faire un meilleur choix de voisinage dans la construction de l'overlay de distribution. Plusieurs solutions nécessitent des modifications pour intégrer l'accès à des informations échangées avec les opérateurs de réseau [AFS07] ou, dans une approche distribuée, juste échangées entre les pairs [CB08]. Des systèmes complets, tels P4P [XYK<sup>+</sup>08], ont été proposés jusqu'à la création d'un groupe de travail à l'IETF : ALTO (*Application Layer Traffic Optimization*) [SKS09, APY<sup>+</sup>14].

Le principal défaut de ces propositions est la nécessité de modifier les applications existantes. Nous nous sommes orientés vers des solutions modifiant les caractéristiques du trafic avec les applications existantes non modifiées. Ces propositions reposent, dans un premier temps, sur une identification des échanges P2P-TV par DPI (*Deep Packet Inspection*) ou par étude comportementale [VR11], puis par une modification des paramètres des trafics en fonction de la localisation des pairs comme présenté sur la figure 1.12.

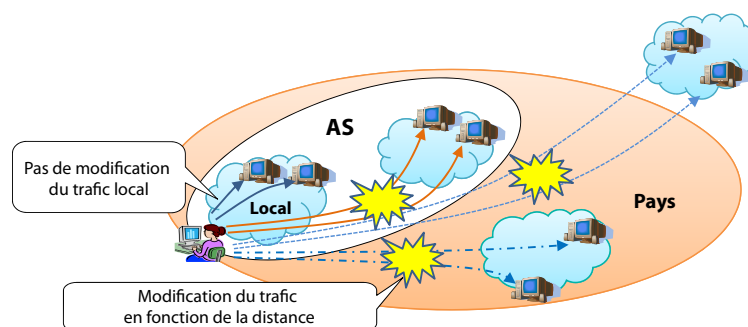


Figure 1.12 – Classification et modification géographique du trafic.

Notre contribution initiale, P2P-DISTO, est une optimisation du trafic P2P par insertion de délais supplémentaires pour les paquets en fonction des emplacements géographiques des pairs [SMF12]. Aucune modification des applications n'est requise : P2P-DISTO est implémenté

sur un routeur d'accès et contrôle les RTT des paquets P2P indépendamment des systèmes P2P. Les mesures expérimentales sur SOPcast et PPStream ont démontré que P2P-DISTO, avec une simple insertion de délais pour les pairs dans un autre pays, permettait de conserver un voisinage de pair similaire, mais relocalisait fortement les trafics de données même lorsque peu de pairs existaient dans la même zone. Par contre, nous avons utilisé une simple correspondance adresse IP / pays avec un délai constant. Nous avons également réalisé un travail complémentaire utilisant la perte de paquet au niveau du routeur [HVMMF13] et observé des résultats similaires.

Pour affiner notre proposition, nous avons ensuite développé P2P-HDISTO, une approche toujours associée au routeur d'accès pour la localisation hiérarchique du trafic P2P [HVSFM14]. Nous avons introduit la distance logique entre deux pairs basée non seulement sur la distance physique, mais aussi sur l'appartenance au même AS, au même opérateur, au même pays, et pour finir au nombre de pairs existants dans la même zone que le pair concerné. Un délai supplémentaire adapté est inséré dans les paquets P2P en fonction de cette distance logique. Des expériences sur les trois applications SOPCast, PPStream et PPTV ont montré une forte réduction des trafics inter-domaines même lorsque peu de pairs existent dans la même zone tout en maintenant les performances des applications (par exemple, le délai de visualisation avec SOPCast passe de 12,3 à 13,1 secondes).

Toujours avec une localisation hiérarchique à plusieurs niveaux, nous limitons cette fois le débit pour les flux associés aux pairs contrôlés [HVMF14]. Comme pour les dégradations précédentes, la baisse de performance de certains échanges améliore la localisation du trafic, ici avec une fonction qui se met en œuvre facilement avec les mécanismes de mise en forme du trafic intégrés dans la plupart des routeurs.

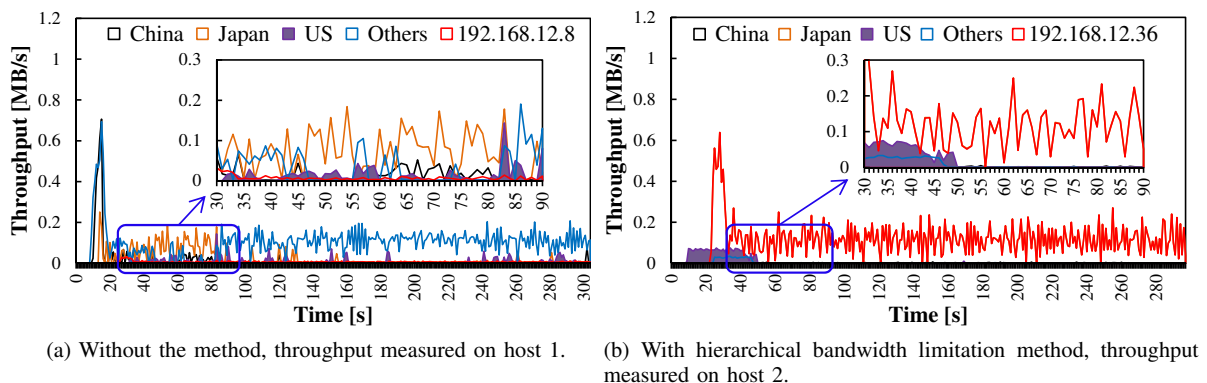


Figure 1.13 – Évolution du débit par région pour SOPCast.

La figure 1.13 montre un exemple de localisation du trafic vidéo en conservant le comportement original de l'application SopCast sur un premier pair (a) et en appliquant la limitation de débit hiérarchique sur un second pair du même LAN (b). Avec la méthode proposée, presque tout le trafic mesuré sur le second pair est téléchargé à partir du premier pair (adresse IP 192.168.12.36). Le temps moyen de démarrage reste stable (12,3 secondes sans contrôle, 12,5 secondes avec notre contrôle de débit hiérarchique). La figure 1.14 présente dans la partie (a) les distributions moyennes des données téléchargées avec trois limitations du débit différentes (sans limite, limitation à 800kb/s hors AS local, limitation hiérarchique). L'axe vertical représente les ratios zone par zone pour le trafic échangé sur le pair supervisé. Nous avons identifié l'AS local (AS4713 NTT), un second FAI japonais très présent dans nos traces (AS SO-Net), les autres AS japonais, la Chine, les États-Unis et les autres pays. Le trafic inter-domaine est fortement réduit, passant de 95% à 20% avec notre proposition. La partie (b) illustre les distributions des pairs voisins

## 1 Distribution vidéo pair-à-pair (P2P-TV)

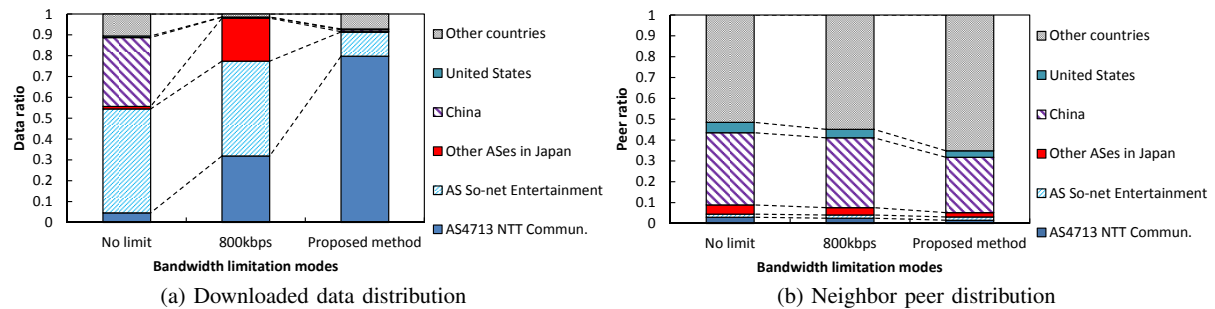


Figure 1.14 – Distribution des données et pairs pour SOPCast avec 3 limitations de débit.

dans les trois modes de limitation de bande passante indiqués précédemment. Les distributions des pairs voisins sont assez stables et semblent relativement indépendantes de la limitation de la bande passante.

## 1.6 Conclusion du chapitre

Les systèmes de P2P-TV sont pour la plupart commerciaux. Ce sont les outils de diffuseurs de contenus vidéo qui doivent rémunérer les chaînes de télévision relayées ou les producteurs de contenus spécifiques. Ces systèmes reposent sur des applications fermées. L'absence d'information sur leurs mécanismes internes fait peser le risque de ne pas contrôler leur trafic sur les opérateurs de réseaux. Nos contributions permettent d'avancer dans la compréhension de ces applications et apportent différentes informations sur la nature des trafics échangés.

Ces nombreuses expériences couvrent la période 2006 à 2016 et ont été réalisées avec un doctorant, Thomas Silverston, puis de nombreux projets de stagiaires de M2 recherche en France ou de MSc au Japon, localisation clé pour ces applications principalement utilisées en Asie.

Les approches utilisées ont été variées, combinant mesures passives d'observation du trafic et mesures actives avec modification du comportement des utilisateurs. Les derniers travaux présentés utilisent des mécanismes de gestion de ressource et introduisent indirectement les travaux du prochain chapitre sur la distribution vidéo via HTTP.

## 2 Distribution vidéo adaptative via HTTP

Au début des années 2010, la distribution de contenu vidéo via P2P est un concept important, particulièrement en Asie. L'évolution des usages et l'amélioration des infrastructures de distribution ont motivé l'adoption d'une autre approche : la distribution adaptative de la vidéo via le protocole HTTP (HAS ou *HTTP Adaptive Streaming*). Celle-ci a supplanté rapidement les autres technologies en dépit de difficultés à surmonter, par exemple la superposition de plusieurs niveaux de régulation du trafic (application, TCP, réseau d'accès...). Nous nous sommes donc intéressés à cette approche et présentons nos contributions dans ce chapitre.

Au cours des années 2010, la distribution de contenu vidéo a généré une part de plus en plus importante du trafic Internet. Le rapport annuel de CISCO indiquait en 2017 [Cisb] que ce trafic avait atteint 70% du trafic Internet total et prévoyait que celui-ci atteindrait plus de 80% du trafic d'ici 2022. Cette énorme demande de transmission de données vidéo sur Internet nous a motivés pour étudier les performances de la distribution vidéo dans les réseaux, et en particulier en bordure où le débit peut être limité et fluctuant. Notre but fut de proposer des solutions pour améliorer la qualité d'expérience de l'utilisateur final dans le cadre de ces nouvelles techniques de distribution de la vidéo.

Les services de distribution vidéo les plus utilisés (Netflix, YouTube...) sont de type OTT (*Over-the-Top*), et reposent donc sur un service réseau non contrôlé. La vidéo est encodée dans les serveurs OTT et livrée à l'appareil de l'utilisateur final via le protocole HTTP. La distribution de la vidéo repose ainsi sur TCP comme couche de transport. Initialement, le téléchargement progressif via HTTP était utilisé et la vidéo était diffusée en utilisant un débit unique. Puis la distribution vidéo adaptative (*adaptive bit rate streaming*) a été introduite pour s'adapter à la bande passante variable disponible entre l'utilisateur final et le serveur OTT. Le lecteur vidéo peut être une application dédiée ou intégrée dans un navigateur Internet. Différents environnements HAS sont proposés [QZS16] tels que : HLS (*HTTP Live Streaming*) développé dès 2009 par Apple pour prendre en charge la distribution de la vidéo sur ses appareils ; SS (*Smooth Streaming*) proposé par Microsoft ; HDS (*HTTP Dynamic Streaming*) proposé par Adobe et s'intégrant dans l'ancienne application de référence pour la vidéo, le lecteur Flash [Ado] ; et DASH (*Dynamic Adaptive Video Streaming over HTTP*), un standard ouvert du consortium MPEG permettant, par exemple, d'utiliser un simple lecteur JavaScript dans un navigateur [AABD12]. Dans la suite, nous avons utilisé DASH comme solution de référence.

Les travaux présentés dans ce chapitre ont été réalisés dans le cadre d'une collaboration avec Anne Fladenmuller pour l'encadrement de la thèse de Rabee Mustapha Abuteir de 2014 à 2018. Les contributions suivantes ont été réalisées : une proposition avec contrôle centralisé via SDN pour améliorer la distribution vidéo dans les réseaux résidentiels [AFF16b] ; son extension pour le sans-fil [AFF16a] ; pour les réseaux d'entreprise [AFF18] ; et un complément sur l'adaptation du débit dans le cadre des réseaux sans-fil avec le professeur Mostafa Ammar de Georgia Tech [AFFA17].

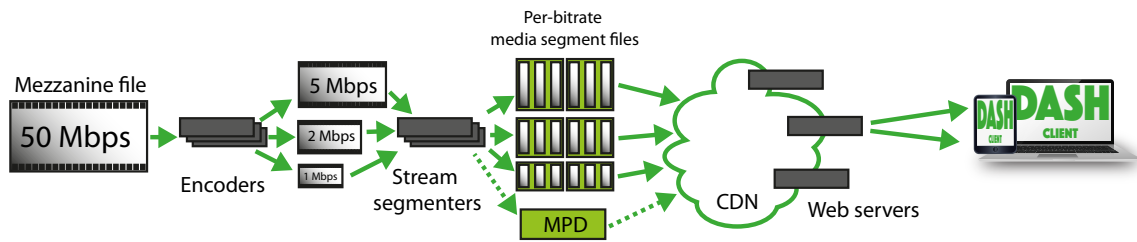


Figure 2.1 – Architecture de bout en bout du système de distribution vidéo DASH

## 2.1 Dynamic Adaptive Video Streaming over HTTP (MPEG-DASH)

DASH a été développé pour la distribution de la vidéo avec une QoE (Qualité d'Expérience) adaptée à chaque utilisateur : le flux vidéo est disponible en plusieurs qualités sur des serveurs Web. La figure 2.1 représente l'architecture de bout en bout classique de la distribution vidéo basée sur la norme MPEG-DASH : la vidéo originale (*Mezzanine file*) est pré-encodée à différents débits et segmentée en morceaux (*chunks*). La norme MPEG-DASH ne standardise que le format de ces segments ainsi que le fichier descriptif MPD (*Media Presentation Description*). Ces données seront ensuite distribuées aux différents serveurs Web impliqués (directement ou via CDN). Le lecteur DASH est un simple client HTTP associé à un algorithme optimisant le choix des segments et dont le développement est laissé libre.

Un même contenu vidéo est ainsi encodé à différents débits puis segmenté en *chunks*, généralement d'une durée de 2 à 10 secondes chacun (en 2016, Netflix utilisait des blocs de 4 secondes avec un débit variable [HJM<sup>+</sup>14]). Au début de l'échange avec le serveur vidéo, le lecteur DASH reçoit les informations sur les débits disponibles parmi d'autres méta-informations vidéo dans le fichier MPD [Bin15]. Un tampon mémoire est utilisé pour lisser la lecture en minimisant l'effet des retards et de la gigue des paquets. Il est rempli au début de la session de distribution : le lecteur télécharge des *chunks* vidéo et remplit la mémoire tampon (*buffering state*). Plus tard, un motif ON-OFF commence à apparaître. Pendant les périodes ON, le lecteur DASH télécharge de nouveaux morceaux tandis que dans les périodes OFF, il attend que le tampon se vide. La taille de la mémoire tampon du client est généralement fixée en secondes, généralement entre 30 et 240 secondes afin de compenser les variations du réseau.

Le lecteur DASH utilise un algorithme d'adaptation au débit pour choisir la succession de *chunks* avec le débit adéquat à demander au serveur de contenu. De nombreux algorithmes ont été proposés dans la littérature et dans les systèmes commerciaux. Quatre catégories sont identifiables pour choisir le prochain débit :

- Adaptation au tampon mémoire (*Buffer Based Adaptation Logic*) : repose sur la mesure du remplissage de la mémoire tampon du lecteur vidéo [HJM<sup>+</sup>14] ;
- Adaptation au débit (*Bandwidth Based Adaptation Logic*) : repose sur l'estimation de la bande passante disponible à partir des *chunks* téléchargés [JSZ12] ;
- Adaptation au délai (*Delay Based Adaptation Logic*) : repose sur la mesure du délai nécessaire pour télécharger les *chunks* [Ado].
- Adaptation hybride (*Hybrid Adaptation Logic*) : choix basé sur plusieurs métriques telles que les mesures du débit de téléchargement et/ou l'occupation du tampon du lecteur, la prédiction de la bande passante et l'apprentissage automatique...[JSM<sup>+</sup>16, MTAA<sup>+</sup>16, GKM<sup>+</sup>16].

Lorsque deux ou plusieurs clients sont en compétition pour la bande passante [AABD12], cela peut entraîner une inadéquation entre les débits demandés par les lecteurs et résulter en une sous-

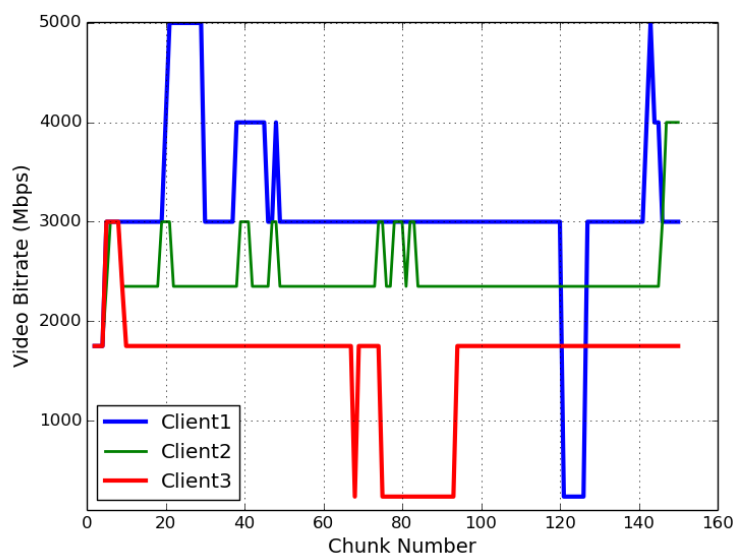


Figure 2.2 – Trois clients recevant simultanément de la vidéo en utilisant dash.js

utilisation de la bande passante du réseau. La figure 2.2 montre dans une simple expérimentation les débits vidéo de trois clients accédant simultanément à du contenu vidéo via DASH sur un réseau filaire. Lorsque le lecteur 1 augmente le débit, cela entraîne une baisse du débit du lecteur 2 et vice versa. La figure montre que les trois flux vidéo ne parviennent pas à partager équitablement la bande passante et que la bande passante totale est sous-utilisée. Bien qu'aucun des clients ne soit affecté par un re-remplissage du tampon mémoire dans cette expérience, les résultats montrent une instabilité et une répartition inéquitable de la bande passante entre les trois flux vidéo.

## 2.2 Distribution vidéo assistée par le réseau : NAVS

Afin de limiter ces instabilités, nous proposons d'introduire une gestion centralisée des ressources à l'aide d'un contrôleur SDN (*Software-Defined Networking*) [ONF12]. Cette technologie permet de découpler le plan de contrôle du plan de données grâce à un contrôleur qui a une vision globale et prend des décisions en temps réel [KF13]. Nous avons ainsi défini une architecture qui repose sur cette approche pour améliorer la distribution vidéo dans les réseaux résidentiels. Cette solution évolutive introduit la possibilité de classification des trafics applicatifs et le contrôle centralisé de l'allocation des ressources [QLJ<sup>+</sup>13, FN14, TCCP15].

L'amélioration de la QoE des utilisateurs dans les réseaux résidentiels est possible avec la mise en forme du trafic HAS [HG12]. Dans cette proposition, cette fonctionnalité de mise en forme (*traffic shaping*) est réalisée dans la passerelle d'accès en fonction du débit souhaité pour chaque flux vidéo. Des approches où le contrôle prend place du côté du serveur ont aussi été proposées [AADB13], mais les causes d'instabilité sont plus facilement identifiables et gérables du côté du lecteur [GCJM12, AMC15], en particulier dans la passerelle d'accès [MFA15]. L'utilisation du SDN a également été abordée, mais avec des modifications du client [GEB<sup>+</sup>13, PWH<sup>+</sup>15].

Les travaux que nous avons réalisés se situent dans un contexte de réseaux résidentiels où plusieurs clients sont connectés au réseau du FAI (Fournisseur d'Accès Internet) via une passerelle d'accès. Dans cette première étude, nous considérons que le goulot d'étranglement du réseau est localisé au niveau du lien d'accès. Lorsque plusieurs flux vidéo sont en compétition pour la bande passante, une mauvaise performance peut en résulter, comme cela a été indiqué précédemment.

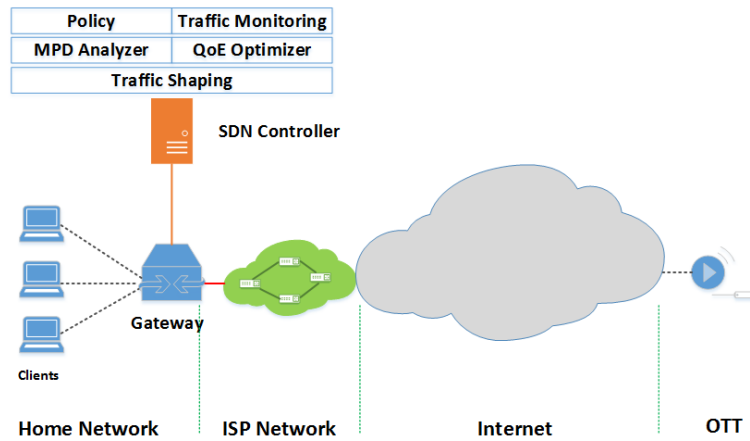


Figure 2.3 – Architecture NAVS pour les réseaux résidentiels

Notre proposition NAVS (*Network Assisted Video Streaming*) est basée sur l'architecture SDN. Elle peut ainsi être implémentée en tant qu'application dans le contrôleur SDN (*Northbound applications*). Elle ne nécessite aucune modification du serveur ou du client, facilitant son intégration dans les systèmes de distribution de la vidéo. NAVS vise à améliorer la QoE des utilisateurs en réduisant l'instabilité des lecteurs, en maximisant l'équité entre les clients et en augmentant la qualité vidéo. NAVS utilise la mise en forme dynamique du trafic vers les clients afin d'atteindre ces objectifs. La figure 2.3 décrit l'architecture proposée avec plusieurs clients connectés à la passerelle d'accès où peut être éventuellement implanté le contrôleur SDN. Cinq modules lui sont associés : le contrôleur de politique, le superviseur de trafic, l'analyseur de MDP DASH, l'optimiseur de QoE et le contrôleur de trafic.

Nous avons expérimenté NAVS par simulation sur NS3 [NS3] avec l'algorithme d'adaptation *Smooth Adaptation* [AABD12, ABD11]. Ces travaux ont montré une amélioration notable en termes d'équité, de stabilité, de QoE et d'utilisation du trafic disponible.

NAVS a été présenté à la conférence IEEE AINA'16 dans l'article « *SDN Based Architecture to Improve Video Streaming in Home Networks* » [AFF16b].

## 2.3 Distribution vidéo assistée pour le sans-fil : WNAVS

L'utilisation de liaisons sans-fil pour accéder à la passerelle dans un réseau résidentiel est de plus en plus répandue, or ce type de technologie peut présenter des inefficacités en termes de débit, en particulier avec TCP [PRR<sup>+</sup>03, HRBSD03], et donc impacter la distribution vidéo dans ce contexte. Plusieurs solutions ont été proposées pour pallier ce problème [LCL11, ROF15, AM15] mais aucune n'est adaptée aux contraintes du trafic vidéo sur TCP. D'autres approches se focalisent sur la QoE [CPZZ15, DHS15, XXCP15] mais les mécanismes proposés s'adaptent mal à la dynamique des applications HAS.

Pour améliorer la QoE dans le contexte des réseaux résidentiels à accès sans-fil, nous avons proposé WNAVS (*Wireless Network Assisted Video Streaming*), qui utilise SDN de manière similaire à notre proposition précédente NAVS. D'autres propositions utilisent SDN [FGB<sup>+</sup>15, MSA<sup>+</sup>15] mais seulement pour la mesure ou dans un contexte plus restreint. Notre système récupère en plus des informations sur les liaisons sans-fil pour optimiser le contrôle du trafic. La figure 2.4 décrit l'architecture proposée avec plusieurs clients connectés en sans-fil à la passerelle avec différentes qualités. La passerelle est connectée à l'Internet via un FAI et est gérée par un contrôleur SDN.

La proposition WNAVS a également été évaluée par simulation (NS3) avec deux algorithmes



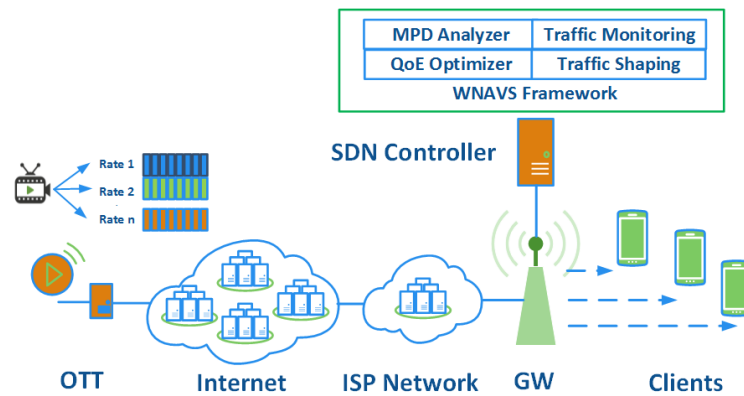


Figure 2.4 – Distribution vidéo dans les réseaux sans-fil résidentiels avec WNAVS

d'adaptation classiques des clients DASH : *Smooth Adaptation* [AABD12, ABD11] et *Buffer Based Adaptation* [HJM<sup>+</sup>14]. Avec du trafic de fond et des clients connectés via des liens aux qualités hétérogènes, WNAVS alloue de la bande passante proportionnellement à la qualité des liens, augmentant ainsi l'utilisation globale de la bande passante et donc la QoE des utilisateurs sans modifier ni le client ni le serveur.

WNAVS a été présenté à la conférence IEEE IWCMC'16 dans l'article « *An SDN Approach to Adaptive Video Streaming in Wireless Home Network* » [AFF16a].

## 2.4 Adaptation aux réseaux sans-fil d'entreprise : SMILE

Les réseaux sans-fil d'entreprise (EWLAN) sont déployés dans de nombreux environnements demandant un accès sans-fil étendu (entreprises, universités, hôtels, centres commerciaux...). Ils consistent en plusieurs points d'accès supervisés par un contrôleur [Cisa, Mer] et interconnectés par des liaisons filaires jusqu'à la passerelle vers le FAI et Internet. Un goulot d'étranglement peut survenir, nous l'avons vu précédemment, au niveau de la passerelle vers le FAI ou d'un point d'accès pour répartir la bande passante entre des clients, mais avec les EWLAN s'ajoute la question de la répartition équitable des ressources entre les différents points d'accès.

Dans ce contexte, nous proposons SMILE, une gestion dynamique de la distribution vidéo dans le cadre des EWLAN à l'aide du SDN. Cette proposition alloue dynamiquement la bande passante entre les clients en fonction de ce qui est disponible, du nombre d'utilisateurs et de la charge des points d'accès. De plus, elle partage équitablement la bande passante entre les points d'accès en fonction de la qualité des liens des clients associés.

La gestion centralisée des EWLAN est un problème ancien [MPC<sup>+</sup>08, SAR<sup>+</sup>09] mais non traité dans le cas de la distribution vidéo. L'utilisation du SDN dans ce contexte a montré son intérêt [SZMC<sup>+</sup>15, SS15, YBC<sup>+</sup>14] et SMILE utilise, par exemple, les fonctions standard prises en charge par OpenFlow et OpenvSwitch (SMILE est compatible avec ces EWLAN utilisant SDN). L'approche centralisée pour assurer l'équité de la qualité vidéo entre les différents clients avait déjà été proposée dans [CDCZ<sup>+</sup>16, JSM<sup>+</sup>16] mais sans tenir compte de la variation de qualité des liaisons sans-fil.

La figure 2.5 montre l'architecture proposée pour gérer la distribution vidéo HAS dans les EWLAN. Les points d'accès et la passerelle sont gérés par un ou plusieurs contrôleurs SDN. Les vidéos sont disponibles sur les serveurs de contenu OTT. La qualité du canal pour chaque lien client-point d'accès peut être hétérogène. SMILE est exécuté sur le contrôleur SDN en tant qu'application afin d'améliorer la QoE des utilisateurs et de lisser la concurrence entre les lecteurs



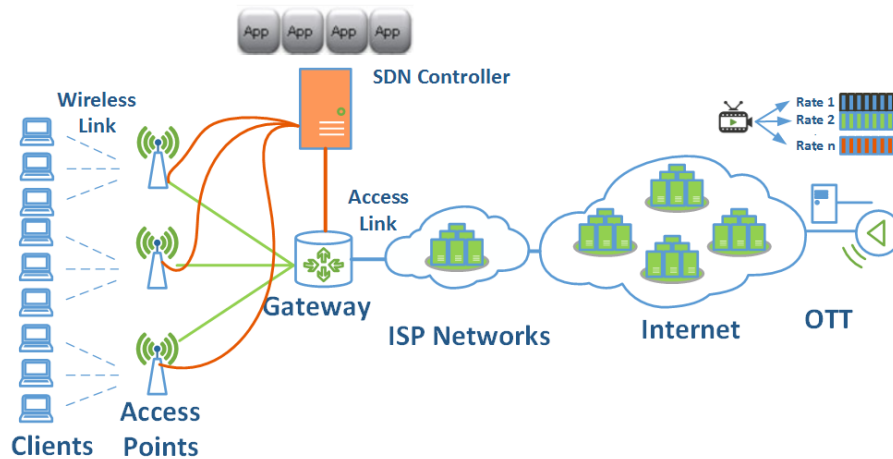


Figure 2.5 – Gestion dynamique et adaptative de la distribution vidéo dans les EWLAN utilisant l’architecture SDN

DASH. La localisation du contrôleur SDN au sein de l’architecture n’est pas précisée (par exemple, il peut se situer dans le contrôleur EWLAN ou dans le *cloud*).

Comme précédemment, nous avons évalué notre proposition via simulation avec NS3 avec les algorithmes d’adaptation usuels. SMILE alloue dynamiquement la bande passante disponible en fonction des charges des points d’accès et la partage équitablement entre les points d’accès en fonction de la qualité des liaisons des clients. SMILE améliore les performances des différents algorithmes d’adaptation DASH sur différentes métriques de performance de distribution vidéo.

SMILE a été présenté à la conférence IEEE PIMRC’18 dans l’article « *Towards Centralized Management for Video Streaming Over Enterprise Wireless LANs* » [AFF18].

## 2.5 Amélioration pour les réseaux mobiles : DAVS

Les approches de distribution adaptative via HTTP utilisent un algorithme d’adaptation au niveau du client pour choisir le débit vidéo pour chaque *chunk* à récupérer depuis le serveur. Pour les clients recevant la vidéo via un réseau cellulaire, le lien cellulaire est généralement le goulot d’étranglement et peut présenter des variations significatives en termes de débit.

En 2017, lors de la visite au LIP6 du professeur Mostafa Ammar de Georgia Tech, nous avons complété nos travaux concernant ces réseaux. Notre contribution porte sur une adaptation dynamique pour la vidéo sur mobiles (*DAVS : Dynamic Adaptation for mobile Video Streaming*), une proposition qui peut être utilisée pour gérer la forte variabilité de débit rencontrée par les clients cellulaires.

L’adaptation au débit (ABR) est utilisée par le lecteur DASH pour le choix des *chunks*. De nombreuses propositions existent [JSZ12, HJM<sup>+</sup>14, Ado] et les schémas d’adaptation typiques reposent sur des décisions basées sur des mesures du débit de téléchargement et/ou de l’occupation du tampon de lecture. L’objectif principal étant de maximiser la satisfaction des utilisateurs en fournissant une vidéo avec la meilleure QoE [DSA<sup>+</sup>11, NKS16]. En général, cela implique une combinaison de maximisation du débit vidéo moyen, de minimisation des événements de blocage du flux vidéo (*rebuffering*) et de minimisation du changement de débit au niveau du client.

Les solutions ABR fonctionnent mieux lorsque la bande passante entre le client et le serveur est relativement stable [MTAA<sup>+</sup>16]. Les performances se dégradent en présence d’une instabilité importante du débit [AFF16a]. C’est généralement le cas pour les clients sur un réseau cellulaire

recevant de la vidéo : des adaptations sont à envisager de manière similaire au contrôle de congestion qui est modifié pour ces réseaux [ZPC<sup>+</sup>15]. Plusieurs propositions ont été faites pour améliorer la situation de la distribution vidéo dans le contexte cellulaire [ZQ16, ZEG<sup>+</sup>15, JTM15, CMK<sup>+</sup>13].

Notre proposition DAVS était spécifiquement conçue pour gérer la forte dynamique non prévisible des réseaux cellulaires. Elle est basée sur la combinaison de deux composants : une gestion dynamique des limites des tampons mémoires (les limites divisent les tampons en zone associée à des politiques de choix) et un composant de stabilisation pour minimiser les changements de débit. L'évaluation a été réalisée par simulation sur des traces réelles de trafic 4G-LTE. Nous avons comparé les performances de DAVS par rapport à quatre algorithmes : OSMF [Ado], BBA2 [HJM<sup>+</sup>14], SARA [JTM15] et PSQA [LL16]. Selon nos simulations, au moment de sa publication, DAVS proposait, pour une stabilité et un *rebuffering* identique aux algorithmes de référence, un meilleur débit moyen et donc une meilleure QoE selon la définition proposée dans [YJSS15].

DAVS a été présenté à la conférence IEEE IWCMC'17 dans l'article « *Variable-Threshold Buffer Based Adaptation for DASH Mobile Video Streaming* » [AFFA17].

## 2.6 Conclusion du chapitre

Les contributions présentées dans ce chapitre sont à mettre en perspective avec la massification des trafics vidéo HAS du milieu des années 2010. Le partage efficace et équitable des ressources n'allait pas de soi dans un environnement où de nombreux éléments sont dynamiques à différentes échelles de temps. Les mécanismes de régulation proposés se sont peu développés, laissant à penser que les algorithmes d'adaptation des clients ont suffisamment progressé pour prendre en compte les différentes situations étudiées.

La portée de ces travaux dépasse les mécanismes d'allocation de ressource proposés pour s'orienter vers les aspects plus généraux de la virtualisation des réseaux. À travers le SDN/NFV (*Network Functions Virtualization*), la chaîne de distribution vidéo évolue avec de nombreux éléments qui se retrouvent actuellement réalisés dans le *cloud*. À courte échéance, ces éléments seront des fonctions virtuelles du réseau ou VNF (*Virtual Network Functions*). Cette évolution permettra une réduction du coût d'infrastructure, ouvrant ces services à des distributions plus légères qui pourront être déployées à la volée.



### 3 Cartographie de l'Internet

L'industrie de la distribution vidéo migre rapidement vers le modèle OTT qui utilise l'Internet pour atteindre l'utilisateur final. Cette approche met en jeu la couche applicative avec une vision de bout en bout des services réseau. Il y a abstraction de la topologie des réseaux utilisés même si des optimisations peuvent être réalisées à l'aide des CDN, par exemple.

Mes travaux antérieurs pour la distribution vidéo se positionnaient au niveau applicatif, que ce soit les approches P2P présentées dans le chapitre 1 ou les propositions reposant sur le HAS (*HTTP Adaptive streaming*) dans le chapitre 2. Ils font complètement abstraction de la topologie du réseau. Dans les deux cas, des informations sur la topologie de l'Internet pourraient les améliorer : avec, par exemple, une localisation des pairs pour construire un meilleur overlay P2P, ou choisir le meilleur serveur Web pour du trafic DASH.

Fin 2016, j'ai eu l'opportunité de démarrer une activité sur la cartographie de l'Internet, une thématique qui apporte un nouvel éclairage sur mes travaux de mesure de la distribution de la vidéo dans l'Internet.

Ces travaux sur la cartographie ont été menés dans le cadre du groupe de recherche que nous supervisons avec Timur Friedman, et plus particulièrement dans le cadre du projet « Cartographie de l'Internet » financé par le ministère des Armées. Ce projet est le support du travail de doctorat de Kévin Vermeulen, qui est au cœur des contributions présentées dans la suite.

Un premier travail a consisté à optimiser l'outil MDA Paris Traceroute, créé dans le cadre d'une collaboration entre le LIP6 et le LIAFA en 2006. Cet outil fournit les routes les plus précises et complètes, mais avec un surcoût en nombre de messages. Nous avons travaillé avec Stephen Strowes de RIPE-NCC pour réduire son coût et essayer de l'intégrer, par exemple, dans les équipements de mesure légers de la plateforme RIPE Atlas [RIP15]. Nous avons publié ces travaux dans la conférence *ACM Internet Measurement Conference (IMC) 2018* [VSFF18].

Pour passer à l'échelle de l'Internet, nous nous sommes ensuite intéressés aux mesures à haute fréquence réalisées par des outils tels que Yarrp. Nous avons ainsi collaboré avec Robert Beverly et Justin Rohrer de NPS (*Naval Postgraduate School*) et proposé l'outil Diamond Miner qui est le plus complet et le plus rapide pour cartographier l'Internet. Ce travail a été publié au *USENIX Symposium on Networked Systems Design and Implementation (NSDI) 2020* [VRB<sup>+</sup>20].

Pour compléter ces travaux, nous avons également collaboré avec le professeur Reza Rejaie de l'université d'Oregon sur la résolution d'alias, une technique qui permet de regrouper plusieurs adresses découvertes appartenant au même routeur. Nous avons publié une nouvelle approche basée sur la limitation en débit des sondes ICMP dans la conférence *Passive and Active Measurement (PAM) 2020* [VLA<sup>+</sup>20].

Ces réalisations n'auraient pas été possibles sans la participation de notre équipe de soutien sur les plateformes PlanetLab Europe et EdgeNet : Burim Ljuma, Vamsi Addanki et Matthieu Gouel. Une publication plus spécifique à l'utilisation de ces plateformes a également été acceptée à *CNERT (Computer and Networking Experimental Research using Testbeds)*, un workshop de IEEE INFOCOM 2019 [VLF<sup>+</sup>19].

### 3.1 Mesure de la topologie de l'Internet : Traceroute

La cartographie de l'Internet est une tâche complexe : de nombreux travaux ont essayé de révéler la topologie de l'Internet [CDZ97, ALWD05, DF07] mais aucun n'a pu produire une vision complète de celle-ci [RWM<sup>+</sup>11, LDcM11].

Plusieurs niveaux d'analyse sont possibles : adresses IP, routeurs (comportant plusieurs interfaces associées à des adresses IP différentes), POP (*Point of Presence*, regroupant plusieurs routeurs co-localisés), AS (*Autonomous Systems*, ensemble de POP et de points d'interconnexion administrativement indépendants qui, ensemble, composent l'Internet).

Si les informations échangées à haut niveau, par exemple entre les AS, sont en partie publiques et permettent d'obtenir simplement une carte des interconnexions à granularité large, la création de cartes plus précises, avec la localisation des routeurs et des réseaux qu'ils interconnectent, est bien plus complexe. L'outil de base pour obtenir des informations à granularité fine (adresse IP) est Traceroute [Jac88], créé en 1988 par Van Jacobson au LBNL à UCB (*University of California, Berkeley*). Cet outil permet de résoudre les problèmes de routage en découvrant les séquences d'adresses IP empruntées à partir d'un point de mesure vers une destination, révélant ainsi la route correspondante. Son utilisation vers plusieurs destinations permet de construire des cartes de l'Internet détaillées [LHH08].

#### Traceroute

Traceroute prend comme paramètre une adresse IP destination. Il émet successivement des messages sonde à destination de cette adresse dans des paquets IP en incrémentant la valeur du champ TTL initial à partir de 1 jusqu'à un événement terminal. Cette valeur de TTL positionné dans le paquet émis est décrémentée dans chaque routeur traversé. Le passage à zéro éventuel du TTL provoque la destruction du message et déclenche l'émission d'un message ICMP « Time Exceed » [Pos81a] vers l'émetteur initial de la sonde avec comme adresse IP source l'une des interfaces du routeur en question.

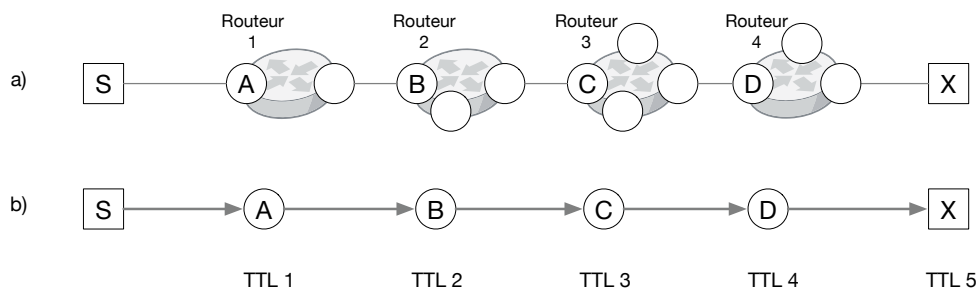


Figure 3.1 – Route de S à X avec (a) les routeurs traversés et (b) la séquence d'interfaces retournée par l'outil Traceroute

L'outil Traceroute retourne ainsi une séquence d'adresses d'interfaces associées à chacun des routeurs successivement traversés (généralement celle d'arrivée dans le routeur) et permet de découvrir la route empruntée à partir d'un point de mesure vers une destination. La figure 3.1 illustre ce mécanisme et présente la découverte de la route S, A, B, C, D, X habituellement représentée sous forme d'un graphe orienté où les nœuds sont identifiés par les interfaces découvertes et sont reliées par des liens orientés dans la direction de la destination. Malheureusement, un certain nombre de problèmes compliquent ce travail, par exemple : les routeurs qui ne génèrent pas de message ICMP ; ceux qui répondent par une interface différente de celle empruntée par la sonde ;

l'existence de chemins asymétriques ; le traitement spécifique des messages ICMP ; l'adressage privé ; l'utilisation de tunnels ; et la présence de répartiteurs de charge (*load balancers*).

Parmi les différentes difficultés à obtenir une route correcte avec Traceroute, la présence de la répartition de charges est la plus pénalisante. Cette technique génère des informations qui peuvent être interprétées de manière incorrecte alors qu'elle est de plus en plus utilisée : en 2019, nous avons mesuré sa présence dans 65% des routes de l'Internet [VRB<sup>+</sup>20].

La répartition de charge est une fonctionnalité des routeurs qui permet de distribuer le trafic vers une même destination à travers plusieurs routes de coût équivalent. Cela permet d'augmenter le débit agrégé et la résistance aux pannes. Deux types de répartition de charge sont configurables sur les routeurs : « déterministe » et « non-déterministe » [Cis07, Jun, Hua19]. Dans le cas déterministe, une fonction de hachage est calculée sur certains champs de l'entête du paquet pour l'envoyer vers l'une des interfaces de sortie empruntant l'une des routes de coût égal. Dans le cas non-déterministe, aussi appelé « par paquets », les paquets sont associés aux interfaces de sortie généralement selon le principe du tourniquet (*round-robin*).

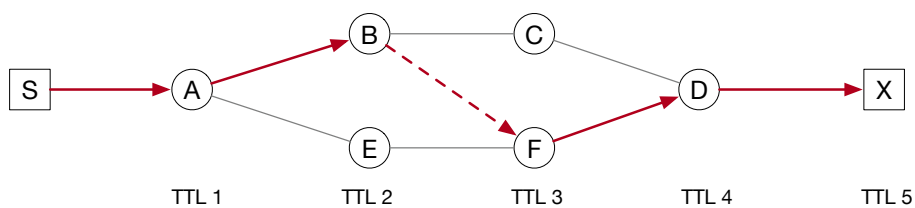


Figure 3.2 – Route de S à X avec une répartition de charge déterministe. En rouge, la succession d'interfaces A, B, F, D, X annoncée par Traceroute avec en pointillé un lien physique inexistant.

La plupart des routeurs utilisent la répartition de charge de type « déterministe » [AFT11, VSFF18]. La combinaison des champs impliqués dans le calcul de la fonction de hachage (un sous-ensemble des champs d'entête tels que l'adresse source, l'adresse destination, le protocole, les ports...) est appelée identifiant de flot. Deux paquets du même flot suivent la même route afin de ne pas trop impacter les performances des couches supérieures. La répartition de charge étant peu présente lors de la création de l'outil Traceroute, celui-ci fait varier les identifiants de flot en incrémentant le port destination (pour apparier les sondes et les réponses). Un biais est ainsi introduit lors du calcul de route : dans la figure 3.2, on observe la route S, A, B, F, D, X qui n'existe pas, mais semble découverte à travers la séquence d'interfaces retournée par Traceroute. La sonde qui a expiré à TTL 2 en B avait l'identifiant de flot de la route du haut et celle qui a expiré à TTL 3 en F avait un identifiant associé à la route du bas.

## Paris Traceroute

Paris Traceroute a été développé pour éviter le problème des sondes successives qui utilisent des identifiants de flot différents [ACO<sup>+</sup>06]. Cet outil fixe l'identifiant d'une suite de sondes afin de garantir le passage par les mêmes interfaces lorsque des routeurs avec répartition de charge déterministe sont présents. La figure 3.3 illustre le passage de toutes les sondes par la même route S, A, B, C, D, X grâce à un identifiant de flot identique (A est l'interface d'un routeur avec répartition de charge déterministe).

Lors du passage par un routeur avec répartition de charge déterministe, Paris Traceroute permet de découvrir une route. Si l'on relance plusieurs fois l'outil, d'autres routes seront éventuellement découvertes avec un identifiant de flot différent, mais ceci sans garantie de les découvrir toutes.

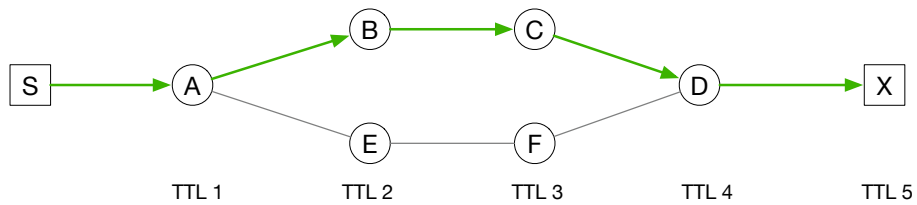


Figure 3.3 – Route de S à X avec une répartition de charge déterministe. En vert, la succession d'interfaces A, B, C, D, X annoncée par Paris Traceroute suivant un même chemin de liens physiques.

### MDA Paris Traceroute

L'algorithme MDA (*Multipath Detection Algorithm*) a été développé pour définir le nombre de sondes minimum à envoyer pour découvrir toutes les routes passant par un routeur avec des garanties statistiques. Pour chaque routeur avec  $k$  liens de sortie en répartition de charge, le nombre de sondes avec des identifiants différents à envoyer pour garantir qu'il n'y en a pas plus de  $k$  est  $n_k$  (sous certaines hypothèses) [VATF09]. Cette valeur  $n_k$ , appelée condition d'arrêt (*stopping point*), indique donc le nombre de sondes à envoyer au TTL suivant d'un routeur dont on souhaite déterminer le nombre de liens impliqués dans une répartition de charge. Par exemple, avec une probabilité d'échec d'énumération de tous les liens suivants d'un routeur de 5%, on a :

$$n_1 = 6, n_2 = 11, n_3 = 16, n_4 = 21, n_5 = 27, \dots n_{10} = 57, \dots n_{100} = 757, \dots$$

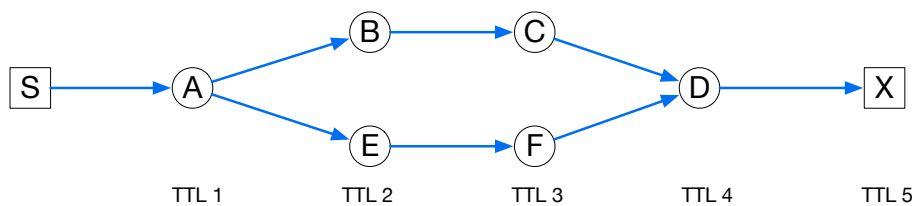


Figure 3.4 – Route de S à X avec une répartition de charge déterministe. En bleu, les successions d'interfaces A, B, C, D, X et A, E, F, D, X découvertes par l'algorithme MDA avec Paris Traceroute.

La figure 3.4 reprend l'exemple précédent, mais en ayant détecté les deux routes possibles vers X. Pour cela, avec une probabilité d'échec de 5%, il a été nécessaire d'envoyer initialement 6 sondes à TTL 2 pour découvrir les deux liens suivant A, puis 5 supplémentaires pour un total de 11 sondes sans découvrir de nouveau lien, pour garantir statistiquement l'absence d'un troisième lien. Les interfaces B, E, C, F, D nécessitent chacune 6 sondes pour garantir qu'elles n'ont qu'un successeur.

Les topologies qu'introduisent les répartiteurs de charge sont appelées des diamants. Ce sont des sous-graphes délimités par un point de divergence suivi, deux ou plus sauts plus loin, par un point de convergence avec la contrainte que tous les flots de la source à la destination passent par ces deux points. Par exemple, la figure 3.4 présente un diamant dont A est le point de divergence et D celui de convergence.

### 3.2 Amélioration de MDA Paris Traceroute : MDA-Lite

MDA Paris Traceroute a été utilisé plus d'un milliard de fois dans le cadre de *Measurement-Lab* [CB18], démontrant ainsi son utilité. Malheureusement, l'algorithme MDA nécessite l'envoi d'un nombre trop important de sondes pour des équipements légers comme les boîtiers RIPE Atlas qui n'utilisent que Paris Traceroute (sans MDA). Notre première contribution résulte d'une collaboration avec RIPE-NCC visant à réduire le nombre de sondes envoyées en conservant les garanties statistiques avec la proposition MDA-Lite.

Reprenons les hypothèses associées à l'algorithme MDA : (1) routage stable pendant le processus de découverte ; (2) répartition de charge de type déterministe ; (3) choix des liens aléatoire et uniforme ; (4) toutes les sondes reçoivent des réponses ; (5) et les répartiteurs de charge sont indépendants.

Si l'on considère une représentation des routes vers la destination sous forme d'un graphe orienté où les nœuds sont identifiés par les interfaces découvertes, l'algorithme MDA se déroule de la manière suivante : tant que l'ensemble des nœuds déjà découverts, mais non encore résolus (c.-à-d. dont tous leurs successeurs n'ont pas encore été découverts), n'est pas vide, un nœud est extrait de celui-ci et est résolu. La résolution d'un nœud  $N$  situé au TTL  $i$  considère que ce nœud a  $k$  successeurs (en démarrant par  $k = 1$ ) et nécessite donc l'envoi de  $n_k$  sondes au TTL  $i+1$  avec des identifiants de flot différents. Les successeurs ainsi découverts sont enregistrés dans l'ensemble des nœuds à évaluer (sauf s'ils l'ont déjà été) et si seulement  $k$  sont observés, alors la condition d'arrêt est atteinte et le nœud est considéré résolu, sinon  $k' > k$  nœuds ont été découverts et  $n_{k'} - n_k$  sondes complémentaires sont envoyées pour essayer d'atteindre la condition d'arrêt avec  $n_{k'}$ .

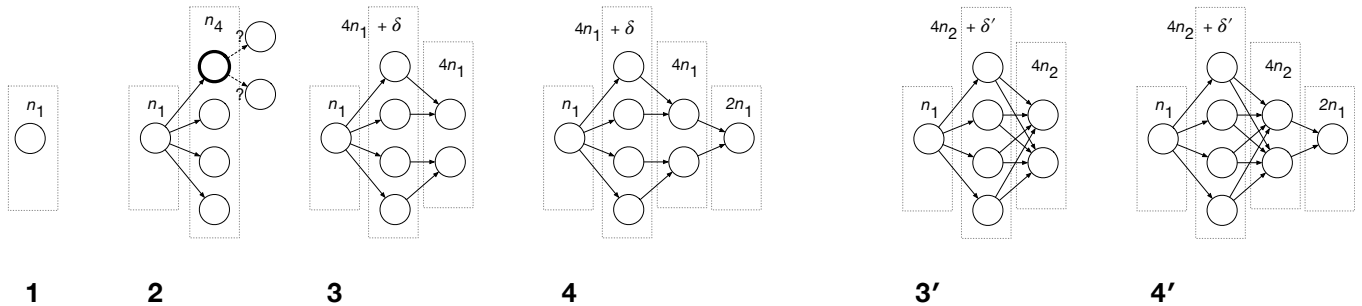


Figure 3.5 – Processus de découverte d'un diamant non-maillé ou maillé avec MDA.

La figure 3.5 illustre le processus de découverte de l'algorithme MDA pour deux types de diamant complètement révélés aux étapes 4 et 4'. Seuls les points de divergence, de convergence, et les sauts intermédiaires sont représentés. Le premier est de type « non-maillé » et le second de type « maillé ». Les conditions d'arrêt sont calculées pour une probabilité d'échec de 5%.

#### Diamant non-maillé

À l'étape 1,  $n_1$  sondes avec des identifiants de flot différents ont été envoyées au saut 1 et ont découvert seulement le point de divergence. À l'étape 2,  $n_1$  sondes sont initialement envoyées au saut 2 et plus d'un successeur sont découverts. La condition d'arrêt évolue vers  $n_2$ , puis  $n_3$  lorsqu'un troisième successeur est découvert, puis  $n_4$  au quatrième. Aucun nouveau successeur n'est alors découvert lorsque cette limite est atteinte. À la prochaine étape, il faut considérer que chacun des nœuds découverts peut avoir plusieurs successeurs, donc garantir que les conditions d'arrêt seront vérifiées pour chacun en garantissant que les sondes adéquates vont bien passer



par les nœuds antérieurs corrects. Ce mécanisme est appelé contrôle des nœuds (*node control*) dans MDA.

Pour le diamant « non-maillé », à l'étape 3, il est nécessaire d'envoyer  $n_1$  sondes pour chaque nœud identifié à l'étape 2. Ce nombre peut être plus grand que ce qui avait été envoyé à l'étape précédente, il faut donc pour les sondes supplémentaires de nouveaux identifiants de flot. Dans notre cas, avec  $n_1 = 6$  et  $n_4 = 21$ ,  $4n_1 = 24 > n_4$ , 3 nouvelles sondes devront alors être utilisées en contrôlant par quel nœud du saut 2 elles proviennent. Si la distribution observée ne propose pas  $n_1$  sondes par nœud, il faudra aussi en rajouter jusqu'à vérifier cette condition. C'est pour cela qu'il est finalement indiqué  $4n_1 + \delta$  avec  $\delta \in \mathbb{N}$  pour le saut 2. Pour le saut 3 on obtient bien  $4n_1$  en choisissant les sondes adéquates pour chaque nœud du saut 2, ce qui nous permet de découvrir deux nœuds. Pour terminer, à l'étape 4,  $2n_1$  sondes sont suffisantes pour découvrir le point de convergence au saut 4. Au total,  $66 + \delta$  sondes sont nécessaires ( $11n_1 + \delta$ ).

#### Diamant maillé

Pour le diamant « maillé », l'augmentation du nombre de liens dans la topologie va accroître le nombre de sondes nécessaires. Les étapes 1 et 2 sont identiques et ne sont pas répétées. À l'étape 3',  $n_2$  sondes sont envoyées pour chaque nœud identifié à l'étape 2 puisqu'ils ont chacun deux successeurs. Ce nombre est plus grand que ce qui avait été envoyé à l'étape précédente, il faut donc également des sondes supplémentaires. Dans notre cas, avec  $n_2 = 11$  et  $n_4 = 21$ ,  $4n_2 = 44 > n_4$ , 23 nouvelles sondes devront alors être utilisées en contrôlant par quels nœuds du saut 2 elles proviennent. Si la distribution observée ne propose pas  $n_2$  sondes par nœud, il faudra aussi en rajouter jusqu'à vérifier cette condition. Nous indiquons ainsi  $4n_2 + \delta'$  avec  $\delta' \in \mathbb{N}$  pour le saut 2. Pour le saut 3, on a  $4n_2$  en choisissant les sondes adéquates pour chaque nœud du saut 2, ce qui nous permet de découvrir deux nœuds. Et à l'étape 4',  $2n_1$  sondes sont suffisantes pour découvrir le point de convergence. Au total,  $106 + \delta'$  sondes sont nécessaires ( $8n_2 + 3n_1 + \delta'$ ).

La seconde topologie introduit la notion de maillage dans un diamant. Cette caractéristique est liée au nombre de liens entre des sauts adjacents. Par exemple, pour le saut entre les nœuds découverts à TTL  $i$  et TTL  $i+1$ , celui-ci sera considéré maillé si une des deux conditions suivantes est validée :

- Le nombre de nœuds au saut  $i$  est inférieur ou égal à celui au saut  $i+1$  et le nombre de liens entrants sur au moins un nœuds au saut  $i+1$  est strictement supérieur à un.
- Le nombre de nœuds au saut  $i$  est supérieur ou égal à celui au saut  $i+1$  et le nombre de liens sortants sur au moins un nœuds au saut  $i$  est strictement supérieur à un.

Un diamant maillé comporte au moins un saut maillé. Dans la figure 3.5, le diamant complètement découvert en 4' est maillé à cause du nombre de liens découverts en 3'.

#### MDA-Lite

Notre première contribution, MDA-Lite, est une évolution de MDA qui se focalise sur les topologies les plus couramment observées : les diamants non maillés uniformes. Dans cette configuration, on peut limiter le contrôle des nœuds et considérer l'algorithme par saut et non par nœud. Ainsi, pour le diamant non maillé découvert en 4 de la figure 3.5, tandis que MDA nécessitait  $66 + \delta$  sondes, MDA-lite n'a besoin que de 44 sondes ( $n_4 + n_2 + 2n_1$ ).

Pour vérifier que l'on peut appliquer MDA-Lite, une détection d'un éventuel maillage est réalisée en utilisant un contrôle des nœuds limité aux diamants de plus de deux sauts et une vérification de la topologie (symétrie) pour garantir l'uniformité nécessaire à l'évaluation saut par saut.

L'efficacité de l'algorithme MDA-Lite a été évaluée par simulation sur des topologies réelles connues. La figure 3.6 présente les résultats pour plusieurs types de topologies : uniforme ou

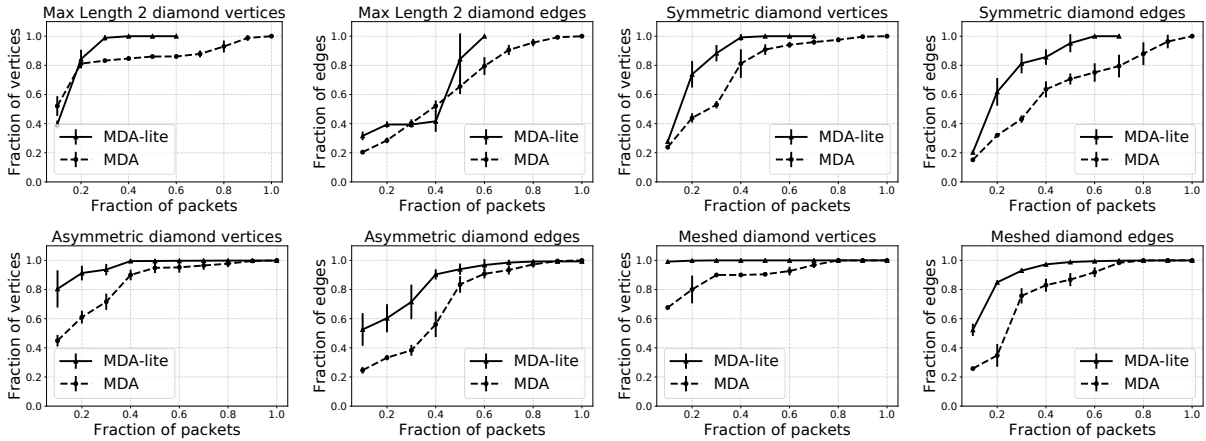


Figure 3.6 – Simulations de MDA-Lite et MDA

asymétrique, maillée ou non (la fraction de paquets envoyés totale est inférieure à 1 lorsque MDA-Lite est dans une situation où l'algorithme nécessite moins de sondes que pour MDA).

Une étude complémentaire basée sur des mesures a également été réalisée. Nous avons utilisé un échantillon de 10 000 routes source-destination comportant des diamants. Les résultats comparatifs moyens obtenus sur ces routes révèlent une économie de l'ordre de 30% des sondes envoyées avec une capacité de découverte similaire :

Paris Traceroute	Nœuds	Liens	Sondes
sans MDA (id. flot unique )	$n$	$l$	$s$
avec MDA	$1,86 \times n$	$4,97 \times l$	$25,0 \times s$
avec MDA-Lite	$1,87 \times n$	$5,00 \times l$	$17,4 \times s$

Évidemment, une grande disparité existe en fonction des configurations : MDA peut envoyer de 2 à 1000 fois plus de sondes pour révéler l'intégralité d'une topologie multi-chemins que Paris Traceroute (sans MDA) pour parcourir une route unique de cette même topologie. MDA-Lite réalise plus de 40% d'économie par rapport à MDA dans 30% des topologies, mais peut aussi être plus coûteux dans 11% des cas.

Au final, si le gain obtenu avec MDA-Lite n'est pas négligeable, il n'est pas suffisant pour intégrer l'outil MDA Paris Traceroute dans des équipements à ressources très limitées comme les sondes RIPE-Atlas. Par contre, l'économie en nombre de sondes réalisée sera appréciable pour les travaux de mesure à haute fréquence que nous présentons dans la section 3.3.

La découverte de la topologie au niveau des adresses IP est souvent la base pour obtenir celle de niveau routeur (en regroupant un certain nombre d'adresses d'interface). Cette résolution d'alias est discutée dans la section 3.4. Certaines techniques utilisées nécessitent de générer à nouveau des sondes pour vérifier l'appartenance de deux ou plusieurs interfaces au même routeur. Nous avons donc associé la résolution d'alias en complément à MDA-Lite dans un nouvel outil appelé MMLTP (*Multilevel MDA-Lite Paris Traceroute*) qui permet, en utilisant les sondes utilisées pour la découverte, de réduire celles nécessaires à la résolution d'alias. MMLTP permet, outre de proposer un outil de type Traceroute avancé avec détection des routeurs et des routes multiples entre une source et une destination, de bénéficier d'un second niveau de réduction du coût global de découverte en nombre de sondes.

Nous avons présenté cette contribution intitulée « *Multilevel MDA-Lite Paris Traceroute* » à *ACM Internet Measurement Conference (IMC'18)* [VSFF18].

### 3.3 Mesure de l'Internet à haute fréquence : D-Miner

Les outils présentés précédemment permettent de découvrir des parties de l'Internet. Leur utilisation systématique à large échelle permet de découvrir la topologie globale d'Internet. Deux systèmes font référence dans ce domaine : Ark et Yarrp.

Ark [CAI19] est le système de cartographie de la topologie Internet développé depuis 2007 par le centre de référence CAIDA (*Center for Applied Internet Data Analysis*) à UCSD (*University of California, San Diego*). Il réalise des mesures en continu à partir de plus d'une centaine de sites distribués géographiquement dans le monde. L'outil de base pour ces mesures est Paris Traceroute utilisé vers une adresse aléatoire dans chaque préfixe /24 annoncé dans le monde. Un ensemble complet de mesures vers tous les préfixes /24 est appelé un cycle et dure en moyenne une journée. Comme l'adresse utilisée dans le /24 est choisie au hasard, l'agrégation des résultats de plusieurs cycles peut révéler des routes multi-chemins, mais ce n'est pas la finalité de ce système qui n'est pas explicitement conçu pour révéler la répartition de charge.

Yarrp [Bev16] est une technique de cartographie de l'Internet à haute fréquence développée par Robert Beverly au NPS (*Naval Postgraduate School*). Il atteint une vitesse élevée en découplant le sondage de la reconstruction de la topologie. En effet, pour l'outil Traceroute et ses évolutions, les réponses sont associées aux sondes correspondantes en faisant varier certaines valeurs des champs des entêtes. En pratique, l'expéditeur maintient une table des sondes envoyées pour pouvoir leur faire correspondre les réponses. Cette approche limite les possibilités de parallélisation. Avec Yarrp, les sondes sont auto-identifiantes : le TTL et l'horodatage auxquels ils ont été envoyés sont codés dans certains champs d'entête qui sont récupérés dans les réponses. De ce fait, on peut reconstruire les séquences Traceroute et la topologie a posteriori. Les approches utilisant le sondage haute fréquence présentent l'inconvénient potentiel d'être limitées en débit par les routeurs. Pour éviter ce problème, Yarrp envoie ses sondes de manière aléatoire pour ne surcharger aucune route.

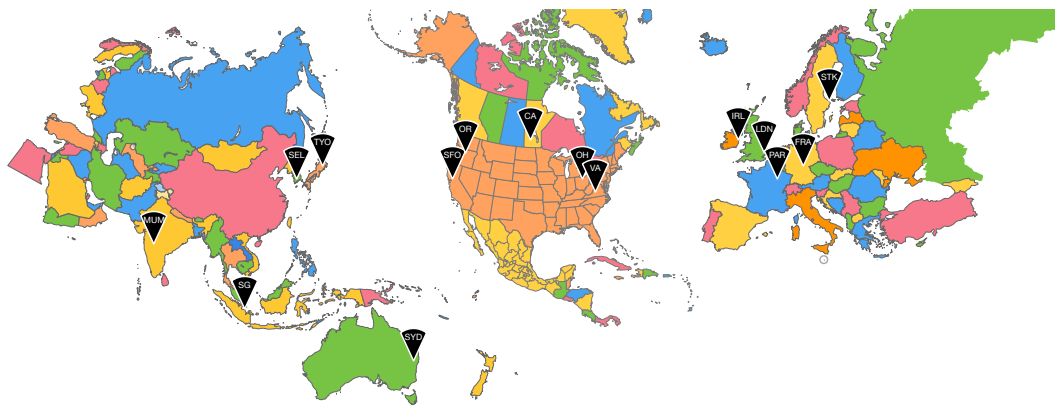


Figure 3.7 – 15 datacenters AWS pour cartographier l'Internet avec Yarrp

Yarrp a été déployé sur le service cloud AWS (*Amazon Web Service*) dans 15 datacenters géographiquement distribués (voir la figure 3.7). Les 15 points de mesures atteignent une fréquence de sondage agrégée de 1 million de paquets par seconde [Roh20]. En sondant une adresse par préfixe /24 comme le système Ark, Yarrp dans AWS est capable d'obtenir des instantanés de topologie au niveau IP d'Internet toutes les 15 minutes. Mais, comme Ark, il n'est pas destiné à découvrir les routes multi-chemins et ne produit qu'une vision partielle.

Pour notre contribution D-Miner (*Diamond Miner*), nous avons collaboré avec Robert Beverly pour développer un outil qui permet la découverte de la topologie de l'Internet à haute fréquence

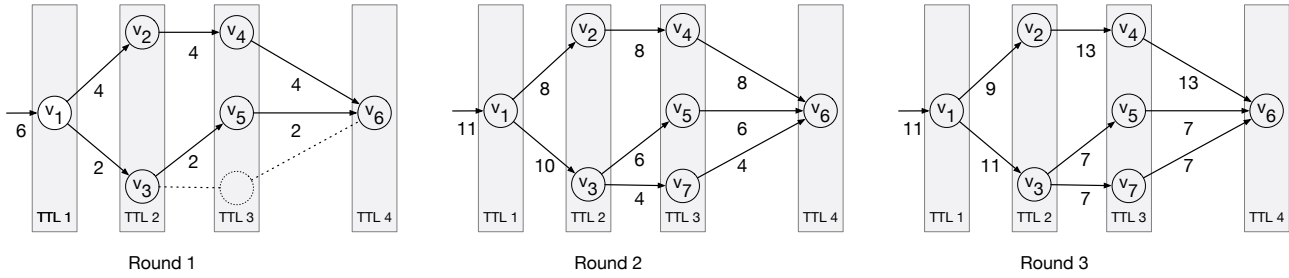


Figure 3.8 – Exemple de topologie découverte en 3 *rounds* par D-Miner (la numérotation des liens indique le nombre de sondes ayant expiré à la prochaine interface rencontrée)

incluant les routes multiples. Cette proposition repose sur l'envoi rapide de sondes de manière aléatoire pour éviter la surcharge d'une seule route, et procède ensuite par itération pour obtenir une vision complète d'Internet avec un maximum de liens et des garanties statistiques.

Les approches de découverte de routes multiples telle MMLPT ou MDA Paris Traceroute effectuent une exploration séquentielle en largeur : à chaque TTL, tous les nœuds sont évalués successivement pour atteindre une probabilité satisfaisante de découverte de tous les liens. Pour D-Miner, le sondage est découplé de l'inférence statistique pour permettre l'envoi aléatoire à haute fréquence des sondes. Le coût total du sondage est amorti en procédant par étapes successives (*rounds*). D-Miner maintient ainsi un ensemble de nœuds non résolus (pour lesquels le nombre de liens est incertain) et procède par *round* jusqu'à ce que la topologie soit statistiquement complète.

La figure 3.8 illustre le fonctionnement de D-miner : soit une source et un préfixe /24 de destination. La topologie présente les découvertes possibles après chacun des 3 *rounds* de sondage. Chaque lien est numéroté avec le nombre de sondes qui ont expiré à l'interface d'entrée du nœud suivant. Au *round* 1, D-Miner a envoyé  $n_1 = 6$  sondes par TTL (les valeurs d'arrêt sont celles utilisées précédemment pour une probabilité d'échec de 5%). Après avoir envoyé 6 sondes et n'avoir découvert qu'un seul successeur, nous avons donc une probabilité de 95% qu'il n'y ait qu'un seul successeur ( $n_1 = 6$ ), alors qu'il faut envoyer 11 sondes pour obtenir la même probabilité qu'il n'y ait que deux successeurs ( $n_2 = 11$ ).

Pour calculer le prochain lot de sondes, en fixant le TTL à  $h$ , nous avons :

- $R_h$  l'ensemble des nœuds  $v_i$  découverts à TTL  $h$ .
- $D_h$  la distribution de probabilité des nœuds répondant pour TTL  $h$  après le *round* de sondage en cours (par exemple, dans la figure 3.8 après le premier *round* de 6 sondes par TTL,  $D_2 = \{v_2 = \frac{4}{6}, v_3 = \frac{2}{6}\}$ ).
- $k_v$  le nombre de successeurs pour le nœud  $v$ .
- $t_h$  le nombre de sondes déjà envoyées à TTL  $h$ .
- $n_k$  le point d'arrêt décrit précédemment.

Étant donné  $h$ ,  $R_h$  et  $D_h$ , nous avons démontré dans [VRB<sup>+</sup>20] que le nombre minimal attendu de sondes nécessaires pour atteindre les garanties statistiques MDA pour tous les éléments de  $R_h$  est :

$$\max_{v \in R_h} \left( \frac{n_{k_v}}{D_h(v)} - t_h \right) \text{ à TTL } h$$

$$\text{et } \max_{v \in R_h} \left( \frac{n_{k_v}}{D_h(v)} - t_{h+1} \right) \text{ à TTL } h + 1$$

Chaque TTL  $h$  génère un certain nombre de sondes supplémentaires pour TTL  $h$  et pour TTL  $h + 1$ . Pour chaque TTL  $h$ , on a donc deux valeurs possibles : celle générée par sondes supplémentaires pour TTL  $h - 1$  et celle des sondes supplémentaires pour TTL  $h$ . Pour que les

conditions requises soient validées dans chaque nœud, le maximum entre ces deux valeurs est choisi. Effectuons l'application numérique sur la figure 3.8.

Après le **round 1**, nous avons découvert la topologie à gauche de la figure :

- Au TTL 1, le nœud  $v_1$  a 2 successeurs,  $D_1(v_1) = 1$ . On a 6 sondes envoyées à TTL 1 et 6 sondes à TTL 2. Le nombre de sondes supplémentaires attendues est  $n_2 - 6 = 5$  pour les TTL 1 et 2.
- Au TTL 2, le nœud  $v_2$  a un successeur et  $D_2(v_2) = \frac{2}{3}$ . Le nœud  $v_3$  a un successeur et  $D_2(v_3) = \frac{1}{3}$ . 6 sondes ont été envoyées à TTL 2 et 6 autres ont été envoyées à TTL 3. Le nombre de sondes supplémentaires attendu est  $3n_1 - 6 = 12$  pour les TTL 2 et 3.
- Au TTL 3,  $v_4$  et  $v_5$  sont similaires à  $v_2$  et  $v_3$ , de sorte que le nombre de sondes supplémentaires attendu est  $3n_1 - 6 = 12$  pour les TTL 3 et 4.

Pour chaque TTL, on prend le nombre maximum de sondes entre TTL et TTL-1. Nous devons ainsi envoyer pour le deuxième *round* : 5 sondes à TTL 1, 12 à TTL 2, 12 à TTL 3 et 12 à TTL 4.

Après le **round 2**, la topologie au centre de la figure est découverte :

- Au TTL 1, nous voyons que  $v_1$  a été résolu, car il a deux successeurs et  $11 = n_2$  flots le traversent.
- À TTL 2,  $v_2$  a également été résolu, car il a un successeur et  $8 \geq n_1$  flots y ont été envoyés.  $v_3$  n'est pas résolu, car maintenant il a deux successeurs et a donc besoin de  $n_2 = 11$  flots qui le traversent. Nous avons  $D_2(v_3) = \frac{10}{18}$  et 18 sondes déjà envoyées aux TTL 2 et 3. Le nombre de sondes supplémentaires attendu est  $\frac{18}{10}n_2 - 18 = \lceil 1.8 \rceil = 2$  pour TTL 2 et 3.
- Au TTL 3, nous voyons que  $v_4$  et  $v_5$  ont été résolus, mais pas  $v_7$ . Nous avons  $D_3(v_7) = \frac{4}{18}$  et 18 sondes déjà envoyées aux TTL 3 et 4. Le nombre de sondes supplémentaires attendu est  $\frac{4}{18}n_1 - 18 = 9$  pour TTL 3 et 4.

Pour chaque TTL, on prend à nouveau le nombre maximum de sondes entre TTL et TTL-1. À la fin, nous devons envoyer pour le troisième *round* : 0 sonde à TTL 1, 2 à TTL 2, 9 à TTL 3 et 9 à TTL 4.

Après le **round 3**, la topologie à droite de la figure est découverte :

Nous voyons que maintenant  $v_3$  et  $v_7$  ont été résolus, de sorte que nous pouvons considérer la topologie entière comme résolue.

Grâce à D-miner, nous avons découvert qu'une grande partie de la topologie d'Internet n'est pas révélée par les systèmes de référence. D-Miner, exécuté depuis six points de mesure de PlanetLab Europe, a découvert plus de 7,1 millions de liens en une semaine, à comparer avec Ark qui pendant la même période, à partir de ses 112 points de mesure, a trouvé 4,4 millions de liens. Yarrp, sondant à partir des mêmes points de mesures que D-miner, a trouvé 2,5 millions de liens. Ces découvertes supplémentaires coûtent 2 à 4 fois plus de sondes par rapport à Ark et Yarrp.

En plus d'évaluer la découverte de nœuds et de liens, nous avons étudié les propriétés algorithmiques et système de D-Miner. La conception itérative par *round* de D-Miner engendre un surcoût de 14% par rapport à MDA Paris Traceroute. Cela s'explique par la conception non séquentielle de D-Miner avec un état restreint à la source (basée sur Yarrp) qui permet des fréquences de sondage très élevées. Nous avons également détaillé les temps d'exécution de D-Miner (dont on peut observer un exemple dans la figure 3.9). Nous avons observé que 10 *round* suffisent pour atteindre des garanties statistiques pour 99% des préfixes de destination /24 et qu'une grande partie du temps est consacré à du calcul ou à des actions système qui pourront être ultérieurement optimisées.

Grâce aux instantanés complets à l'échelle d'Internet collectés avec D-miner, nous pouvons étudier les changements et la dynamique de la topologie de l'Internet. Nous avons par exemple constaté que 52% des événements identifiés « changements de routage » ne le sont pas, mais sont en réalité des ré-indexations internes des répartiteurs de charge (*remapping*).

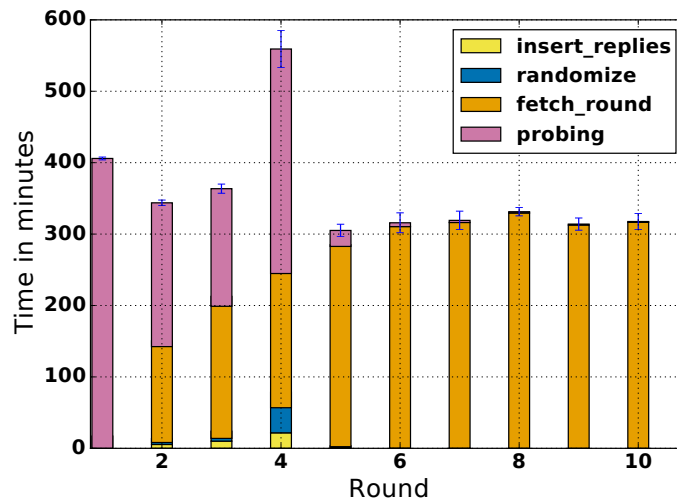


Figure 3.9 – Temps passé dans chaque étape de D-miner pour chaque round (moyenne de trois instantanés) sur une unique machine avec 16 cœurs et 192Go de RAM

Cette contribution a été publiée dans l'article « *Diamond-Miner : Comprehensive Discovery of the Internet's Topology Diamonds* » à la conférence *USENIX Symposium on Networked Systems Design and Implementation (NSDI'20)* [VRB<sup>+</sup>20]. L'outil D-Miner est actuellement opérationnel et nous collectons des instantanés de topologie régulièrement afin de caractériser la prévalence, la taille, l'étendue et l'emplacement de la répartition de charge dans l'Internet actuel.

### 3.4 Cartographie de niveau routeurs : MMLTP et Limited Ltd

Les cartographies évoquées précédemment sont réalisées au niveau des adresses IP, c'est-à-dire au niveau interfaces des équipements interconnectés (principalement des routeurs). Regrouper les adresses IP pour découvrir les routeurs est appelé « résolution d'alias ». C'est un domaine de recherche actif depuis plus de deux décennies. Les techniques utilisées sont basées sur la signature, la topologie ou d'autres informations :

- L'approche basée sur la signature fonctionne en envoyant des paquets de sondes aux différentes adresses IP afin d'obtenir des réponses qui affichent une caractéristique suffisamment distinctive pour permettre d'en identifier certaines comme provenant éventuellement d'un même routeur. Les premiers travaux sur cette approche par Pansiot et Grad utilisaient l'adresse IP source [PG98], concept étendu par l'outil Mercator [GT00]. Plusieurs outils ont utilisé le champ IP ID comme signature : Ally [SMW02], RadarGun [BSS08], et Midar [KHLC13]. Une option « timestamp » peut aussi être utilisée, par exemple pour Pythia [MPP13], et plus spécifiquement pour IPv6, l'entête de fragmentation peut être ciblé, par exemple par l'outil Speedtrap [LBB<sup>+</sup>13].
- L'approche basée sur la topologie fonctionne en envoyant des paquets sondes à des sauts successifs afin d'obtenir des séquences d'adresses IP, puis en raisonnant sur la base de la topologie de niveau IP pour déduire la topologie de niveau du routeur. On peut citer les outils Apar [GS09] et Kapar [Key10] qui se basent sur Traceroute. L'outil DisCarte [SBS08] utilise l'option « Record Route » d'IP.
- Les autres méthodes consistent à consulter d'autres sources, tel le système des noms de domaine. Les outils Rocketfuel [SMW02] et Aroma [KH07] utilisent les informations obtenues grâce aux requêtes *Reverse DNS*.

## MMLTP

Dans un premier temps, nous nous sommes intéressés à la résolution d'adresse pour la combiner avec MDA-lite Paris Traceroute et créer MMLPT (*Multilevel MDA-Lite Paris Traceroute*), un outil efficace en nombre de sondes proposant des informations de niveau routeur des routes parcourues.

Nous avons intégré dans MMLPT les techniques de résolution d'alias existantes, notamment le test des limites de croissance monotone (*MBT, Monotonic Bound Test*) de Midar pour comparer les séries chronologiques des identifiants IP des réponses aux sondes. Le MBT lui-même s'appuie sur les approches pionnières d'Ally et de RadarGun. Nous utilisons également les signatures réseau (*Network Fingerprinting*) de Vanaubel et al. [VPMD13] pour examiner les TTL des réponses aux sondes afin de déterminer quel type de routeur pourrait les avoir envoyées et, combinées à l'examen de la pile d'étiquettes MPLS qui est reçue dans un message ICMP « Time Exceeded », pour tracer les éventuels tunnels MPLS d'un chemin.

Si MMLPT n'apporte aucune contribution à la résolution d'alias en soi (sauf en diminuant le nombre de sondes nécessaires en réutilisant les informations recueillies à partir des sondes envoyées pour découvrir le niveau IP), il facilite l'analyse des routes multi-chemins au niveau des routeurs avec une seule ligne de commande. Avant MMLPT, il fallait d'abord exécuter une technique de découverte de niveau IP, puis transformer la sortie en une entrée correcte pour la technique de résolution d'alias, qui elle-même nécessitait souvent des métadonnées ou une installation supplémentaire [Luc10, KHL13, LBB<sup>+</sup>13].

## Limited Ltd.

Aucun des nombreux outils de résolution d'alias n'offre une technique utilisable en toute circonstance. Midar s'appuie sur les ID IPv4 à croissance monotone, mais de moins en moins de routeurs les utilisent de cette manière. Speedtrap est dédié à IPv6, mais ne peut fonctionner que sur des routeurs avec un identifiant d'entête de fragmentation également à croissance monotone (ce qui exclut certains équipements tels les routeurs Juniper).

Notre contribution à la résolution d'alias s'appelle Limited Ltd. et se prononce « limited limited ». C'est un outil qui ne repose pas sur la croissance monotone d'identifiants, mais propose un nouveau mécanisme qui peut résoudre des alias en IPv4 et IPv6 (incluant les routeurs IPv6 Juniper pour la première fois) et qui continuera à fonctionner sur les routeurs IPv4 pour lesquels Midar ne sera plus adapté.

L'approche classique pour la résolution d'alias consiste à envoyer des sondes à des adresses IP, provoquant des réponses qui affichent une caractéristique suffisamment distinctive pour constituer une signature. Cela permet de faire correspondre celles provenant d'un même routeur. Le nouveau type de signature que nous utilisons est basé sur une fonctionnalité, la limitation de débit ICMP, dans laquelle un nœud (routeur ou hôte final) limite le trafic ICMP qu'il envoie ou reçoit dans un certain laps de temps. Le principal intérêt d'une telle signature est sa large applicabilité pour la résolution d'alias, en particulier pour IPv6 grâce à la limitation du débit ICMP requise pour ce protocole.

Limited Ltd. utilise le mécanisme de ping, requis de tous les hôtes connectés à Internet, par lequel un paquet ICMP « Echo Request » génère en retour un paquet ICMP « Echo Reply ». En envoyant à une adresse IP des pings avec une fréquence assez élevée pour déclencher la limitation des messages ICMP « Echo Reply », tout en envoyant simultanément d'autres pings à d'autres adresses IP à une fréquence réduite, on peut observer si ces bornes limitent également les messages ICMP.

Cette stratégie de sondage produit une trace de perte distincte pour chaque interface. Nous avons constaté que lorsque des pertes se produisent simultanément au niveau de paires d'inter-

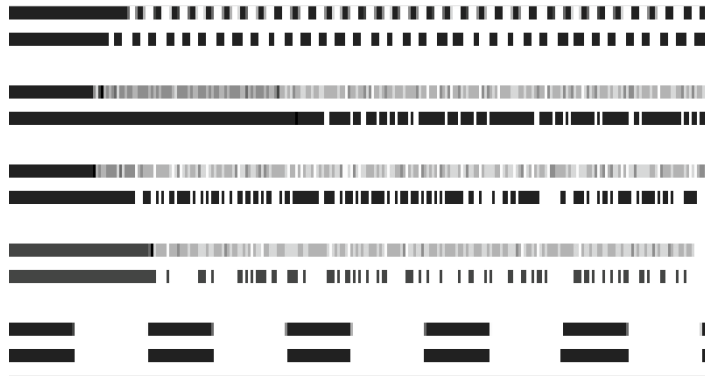


Figure 3.10 – Série temporelle de pertes de paires d'alias

faces alias, elles peuvent le faire de plusieurs manières, comme l'illustrent les cinq exemples de la figure 3.10. Les traits noir et blanc de chaque trace correspondent respectivement aux réponses d'écho ICMP reçues et perdues, et leurs motifs variés ne permettent pas de trouver des corrélations simples. Nous utilisons donc un classificateur d'apprentissage automatique pour identifier des paires d'alias. Il est basé sur les caractéristiques suivantes extraites des traces de perte qui, intuitivement, permettent de capturer le modèle temporel des pertes dans chaque trace : taux de perte, point de changement, probabilités de transition et coefficient de corrélation.

Le ping a tendance à être un mécanisme robuste sur les routeurs modernes, la limitation du taux de réponse d'écho ICMP étant généralement déclenchée à quelques milliers de paquets par seconde. Nous avons vérifié sur des équipements classiques que la surcharge induite était négligeable. De plus, notre approche peut être utilisée de manière parcimonieuse en complément des autres solutions existantes basées aussi sur des signatures.

Nous avons ensuite évalué Limited Ltd. en comparant ses performances à deux outils de résolution d'alias de référence, Midar pour IPv4 et Speedtrap pour IPv6. Nous avons utilisé environ 10 000 adresses IPv4 et IPv6, chacune choisie au hasard parmi les préfixes routables d'Internet. Pour IPv4, notre outil a amélioré la couverture de Midar d'un peu plus d'un quart. Pour IPv6, il a amélioré la couverture de près d'un facteur trois par rapport à l'utilisation de Speedtrap seul. Nous avons étayé cela par des tests sur des topologies réelles dont nous connaissons les cartes exactes fournies par les réseaux Internet2 et Switch.

Nous avons publié « *Alias Resolution Based on ICMP Rate Limiting* » à la conférence *Passive and Active Measurement (PAM'20)* [VLA<sup>+</sup>20].

### 3.5 Conclusion du chapitre

La thématique de la cartographie d'Internet est un sujet fondamental pour comprendre ce réseau qui est devenu notre principal système d'échange d'informations. Que ce soient les utilisateurs qui souhaitent un service performant quelle que soit leur localisation, les opérateurs qui doivent maintenir une infrastructure fiable et sécurisée, ou les régulateurs en charge de vérifier l'équité des conditions d'accès, une connaissance détaillée de la topologie est primordiale.

Nous avons réalisé des contributions reconnues et développé des outils actuellement en exploitation. Nous travaillons à présent sur de nombreuses pistes d'amélioration et d'extension de nos travaux présentés dans le chapitre suivant dédié aux perspectives.





# Perspectives

Nos travaux actuels sont réalisés au sein du groupe de recherche Dioptra où nous développons notre expertise large autour de la mesure et de la cartographie de l'Internet. Trois directions de recherche sont envisagées dans ce contexte.

Un premier axe de recherche fait directement suite aux travaux présentés dans le chapitre 3. Les possibilités de contribution sont importantes, tant au niveau des mesures, de la production de cartes d'Internet et de leur analyse.

En plus des algorithmes et des outils de mesure résultant de nos travaux, nous développons et administrons des plateformes de mesure à large échelle. Ces plateformes utilisent des technologies récentes, par exemple le *Edge Cloud*, et soulèvent des problématiques spécifiques à nos besoins. Ces problématiques seront développées dans un second axe de recherche.

À plus long terme, nous comptons renouer avec nos travaux sur la distribution vidéo en nous basant sur les outils et les plateformes que nous déployons aujourd'hui. Les topologies suivies par les contenus peuvent être très différentes des routes usuelles. Leur exploration par des outils adaptés nous semble fondamentale dans la mesure où la majeure partie du trafic y est associée. Ce sera le troisième axe de recherche.

## Groupe de recherche Dioptra

Les contributions du chapitre 3 et nos recherches actuelles s'intègrent dans le groupe Dioptra que nous animons avec Timur Friedman (<https://dioptra.io>). Ce groupe de recherche associe actuellement deux permanents (Timur Friedman et moi-même), un postdoctorant (Maxime Mouchet) et deux doctorants (Matthieu Gouel et Berat Can Şenel). Nous faisons partie de l'équipe NPA du LIP6 (UMR 7606 SU-CNRS) et travaillons en proche collaboration avec le LINCS (laboratoire commun INRIA, Mines-Telecom, Nokia Bell Labs, SU et SystemX).

Dans le cadre du groupe Dioptra, nous sommes organisés autour de cinq activités : *Dioptra Data* propose l'accès à nos jeux de mesure et à notre système de mesure Iris ; *Dioptra Education* permet de transférer les connaissances vers les étudiants avec des MOOC et des accès spécifiques à nos plateformes de mesures ; *Dioptra Platforms* opère deux plateformes de recherche, EdgeNet et Iris ; *Dioptra Tools* concerne la distribution de nos outils tels que Paris Traceroute ou EdgeNet sous forme de logiciels libres ; et *Dioptra Research* se focalise sur la recherche dans le domaine de la mesure de l'Internet et sur les technologies nécessaires aux outils et plateformes utilisées.

Nos financements actuels reposent principalement sur le ministère des Armées pour la thématique Cartographie de l'Internet. Pour le développement de la plateforme EdgeNet, nous avons reçu des donations de l'entreprise VMware et un financement européen à travers les appels ouverts du projet Horizon 2020 (Fed4FIREplus et NGIatlantic.eu).

À plus long terme, le succès de « SLICES », un projet européen porté par SU, nous laisse envisager de nombreuses opportunités pour nos travaux. « SLICES » est la première grande infrastructure de recherche dédiée aux plateformes d'expérimentation réseau/informatique à large échelle. Elle est financée dans le cadre des projets 2021 de ESFRI (Forum stratégique européen sur les infrastructures de recherche).

## Axe de recherche lié aux mesures de l'Internet

Dans le cadre de la cartographie de l'Internet, nous ne disposons pas encore de la capacité de découverte du système de référence Ark, qui opère à partir de 110 points de mesure. Nous travaillons donc sur l'évolution de nos outils vers un système distribué, Iris (<https://iris/dioptra.io>), qui repose sur des points de mesures hétérogènes : serveurs dédiés très puissants (dans quelques institutions avec lesquelles nous collaborons), machines de puissance intermédiaire (sur notre plateforme EdgeNet), agents de mesure légers (tels des sondes RIPE-Atlas), et éventuellement des ressources dans des *cloud* publics tels AWS (*Amazon Web Services*), GCP (*Google Cloud Platform*) ou Azure.

Ce travail de recherche visant à optimiser la distribution des mesures est mené avec Matthieu Gouel qui débute son doctorat co-encadré par Timur Friedman et moi-même. Maxime Mouchet est également fortement impliqué. Les pistes d'extension de ces travaux sont nombreuses : amélioration des algorithmes pour augmenter la qualité et la vitesse de mesure (actuellement tous les préfixes IPv4 sont sondés en moins d'une journée), enrichissement des informations avec, par exemple, la géolocalisation, l'intégration d'IPv6, la résolution d'alias au niveau routeur et POP.

Nos outils de mesure fonctionnent en permanence, nous permettant d'avoir une vision hebdomadaire de l'Internet et de constituer une base de données pour les études longitudinales de l'Internet disponible pour la communauté scientifique.

## Axe de recherche dédié au Edge Cloud

Pour assurer nos mesures, nous avons développé la plateforme d'expérimentation EdgeNet reposant sur la technologie Kubernetes utilisée à large échelle (<https://www.EdgeNet.org>). Kubernetes a l'avantage d'être un standard industriel, mais n'a pas été conçu pour une approche de type « *edge cloud* » à contributeurs et utilisateurs multiples qui correspond à nos besoins [C $\S$ MC<sup>+</sup>21]. Les problématiques de gestion des accès, de communication à travers les nœuds distants et d'ordonnancement des tâches à large échelle en environnement hétérogène sont non triviales.

Berat Can Şenel a débuté son doctorat sur l'orchestration dans l'*edge cloud* co-encadré par Timur Friedman et moi-même. Nous sommes également assistés par Maxime Mouchet. Ces travaux ont suscité l'intérêt de VMware Research, qui nous finance, et plus récemment du CNCF (*Cloud Native Computing Foundation*) Research User Group. La publication des premiers résultats dans l'article « *EdgeNet : A Multi-Tenant and Multi-Provider Edge Cloud* » dans le *workshop* ACM EdgeSys'21 a reçu le *Best Paper Award* [S $\S$ MC<sup>+</sup>21].

La plateforme EdgeNet est en exploitation avec une cinquantaine de nœuds répartis à travers le monde. Plusieurs autres projets en plus de notre système de mesure Iris et de nos étudiants l'utilisent. En participant à son développement et son administration, nous bénéficions d'une expérience précieuse pour alimenter nos travaux de recherche. Elle sera également le support idéal pour les valider.

## Axe de recherche dédié à la cartographie pour la distribution de la vidéo

CAIDA organise annuellement des groupes de travail réputés autour de la cartographie de l'Internet. Lors de notre participation à AIMS (*Workshop on Active Internet Measurements*), nous avons constaté l'importance des régulateurs (*policymakers*) parmi les participants. Mesurer, analyser et comprendre le fonctionnement de l'Internet n'est pas un domaine réservé aux académiques, aux opérateurs de réseaux ou aux agences de sécurité. Depuis que l'Internet est utilisé comme un moyen de communication essentiel, il devient un service public et doit en tant que tel être régulé afin de garantir l'égalité d'accès aux utilisateurs. Des organismes tels que l'ARCEP

(Autorité de régulation des communications électroniques, des postes et de la distribution de la presse) en France ou la FCC (*Federal Communications Commission*) aux États-Unis se chargent de faire appliquer les politiques d'accès définies par les pouvoirs publics.

Dans le chapitre 2, nous avons indiqué que la distribution vidéo évoluait fortement vers l'approche OTT reposant sur Internet. La diffusion de contenu audio/vidéo est elle-même un domaine traditionnellement régulé. Cette régulation va donc s'appliquer aussi à la distribution de ces contenus via Internet. Il va devenir essentiel de pouvoir mesurer et analyser la distribution de la vidéo.

Le trafic vidéo représente une si grande part du trafic de l'Internet que sa distribution est optimisée et ne suit pas forcément les routes indiquées par les cartes habituelles. Les diffuseurs utilisent des systèmes de réplication qui oriente l'utilisateur vers différentes sources. L'évaluation de la structure de distribution globale n'est pas évidente à inférer de même la QoE pour l'utilisateur final.

Nous proposons donc d'étendre nos outils pour fournir une cartographie de la distribution des contenus et de réaliser des mesures à partir de nos plateformes à large échelle. Développer des techniques spécifiques pour découvrir les infrastructures de distribution n'est pas trivial, car une grande partie de la chaîne de distribution est virtualisée dans le *cloud* et que la distribution finale peut reposer sur des infrastructures multi-CDN dynamiques.

## Conclusion

Ce chapitre de perspectives présente deux axes de recherche déjà engagés et prometteurs à moyen terme. Le troisième et dernier axe de recherche proposé est davantage exploratoire et nous permet de retrouver la thématique de la distribution vidéo, fil conducteur de ce mémoire.



## Bibliographie

- [AABD12] Saamer Akhshabi, Lakshmi Anantakrishnan, Ali C Begen et Constantine Dovrolis : What Happens When HTTP Adaptive Streaming Players Compete for Bandwidth? *In Proceedings of the 22nd international workshop on Network and Operating System Support for Digital Audio and Video*, pages 9–14. ACM, 2012.
- [AADB13] Saamer Akhshabi, Lakshmi Anantakrishnan, Constantine Dovrolis et Ali C Begen : Server-Based Traffic Shaping for Stabilizing Oscillating Adaptive Streaming Players. *In Proceeding of the 23rd ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pages 19–24. ACM, 2013.
- [ABD11] Saamer Akhshabi, Ali C Begen et Constantine Dovrolis : An Experimental Evaluation of Rate-Adaptation Algorithms in Adaptive Streaming over HTTP. *In Proceedings of the second annual ACM conference on Multimedia systems*, pages 157–168. ACM, 2011.
- [ACO<sup>+</sup>06] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien et Renata Teixeira : Avoiding Traceroute Anomalies with Paris Traceroute. *In Proc. IMC '06*, 2006.
- [Ado] Adobe : Open Source Media Framework. [https://www.adobe.com/devnet/video/articles/osmf\\_overview.html](https://www.adobe.com/devnet/video/articles/osmf_overview.html). (Accessed March 06, 2019).
- [AFF16a] Rabee Mustapha Abuteir, Anne Fladenmuller et Olivier Fourmaux : An SDN approach to adaptive Video Streaming in Wireless home networks. *In Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*, pages 321–326. IEEE, 2016.
- [AFF16b] Rabee Mustapha Abuteir, Anne Fladenmuller et Olivier Fourmaux : SDN Based Architecture to Improve Video Streaming in Home Networks. *In The 30th IEEE International Conference on Advanced Information Networking and Applications (AINA-2016)* , Crans-Montana, Switzerland, mars 2016.
- [AFF18] Rabee Mustapha Abuteir, Anne Fladenmuller et Olivier Fourmaux : Towards Centralized Management for Video Streaming Over Enterprise Wireless LANs. *In PIMRC 2018 - IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Bologna, Italy, septembre 2018. IEEE.
- [AFFA17] Rabee Mustapha Abuteir, Anne Fladenmuller, Olivier Fourmaux et Mostafa Ammar : Variable-Threshold Buffer Based Adaptation for DASH Mobile Video Streaming. *In IWCMC 2017 - 13th International Wireless Communications and Mobile Computing Conference*, Valencia, Spain, juin 2017. IEEE.
- [AFS07] Vinay Aggarwal, Anja Feldmann et Christian Scheideler : Can ISPs and P2P Users Cooperate for Improved Performance? *SIGCOMM Comput. Commun. Rev.*, 37(3): 29–40, juillet 2007.
- [AFT07] Brice Augustin, Timur Friedman et Renata Teixeira : Multipath Tracing with Paris Traceroute. *In Proc. E2EMON '07*, 2007.
- [AFT11] Brice Augustin, Timur Friedman et Renata Teixeira : Measuring Multipath Routing in the Internet. *IEEE/ACM Transactions on Networking (TON)*, 19(3):830–840, 2011.

- [AI11] Vetro Anthony et Sodagar Iraj : The MPEG-DASH standard for multimedia streaming over the internet. *IEEE Multimedia*, (4):62–67, 2011.
- [ALWD05] D. Alderson, Lun Li, W. Willinger et J.C. Doyle : Understanding Internet topology : principles, models, and validation. *IEEE/ACM Transactions on Networking*, 13(6): 1205–1218, Dec 2005.
- [AM15] Abdallah S Abdallah et Allen B MacKenzie : A cross-layer controller for adaptive video streaming over IEEE 802.11 networks. *In Communications (ICC), 2015 IEEE International Conference on*, pages 6797–6802. IEEE, 2015.
- [AMC15] Chiheb Ben Ameer, Emmanuel Mory et Bernard Cousin : Evaluation of Gateway-Based Shaping Methods for HTTP Adaptive Streaming. *In 2015 IEEE International Conference on Communication Workshop (ICCW)*, pages 1777–1782. IEEE, 2015.
- [APY+14] R. Alimi, R. Penno, Y. Yang, S. Kiesel, S. Previdi, W. Roome, S. Shalunov et R. Woundy : Application-Layer Traffic Optimization (ALTO) Protocol. RFC 7285, RFC Editor, September 2014. <http://www.rfc-editor.org/rfc/rfc7285.txt>.
- [BBK02] Suman Banerjee, Bobby Bhattacharjee et Christopher Kommareddy : Scalable Application Layer Multicast. *In Proceedings of the 2002 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM '02*, pages 205–217, New York, NY, USA, 2002. Association for Computing Machinery.
- [Bev16] Robert Beverly : Yarrp'ing the Internet : Randomized High-Speed Active Topology Discovery. *In Proc. ACM IMC '16*, 2016.
- [Bin15] Benny Bing : *Next-generation video coding and streaming*. John Wiley & Sons, 2015.
- [BSS08] Adam Bender, Rob Sherwood et Neil Spring : Fixing Ally's Growing Pains with Velocity Modeling. *In Proc. IMC '08*, 2008.
- [CAI19] CAIDA : The CAIDA ark IPv4 Internet topology data kits dataset, 2019. <http://www.impactcybertrust.org>.
- [CB08] David R. Choffnes et Fabián E. Bustamante : Taming the Torrent : A Practical Approach to Reducing Cross-Isp Traffic in Peer-to-Peer Systems. *SIGCOMM Comput. Commun. Rev.*, 38(4):363–374, août 2008.
- [CB18] Ya Chang et Peter Boothe : A better schema for paris-traceroute. Presentation at AIMS '18 CAIDA Workshop, Mar 2018.
- [CDCZ+16] G Cofano, L De Cicco, T Zinner, A Nguyen-Ngoc, P Tran-Gia et S Mascolo : Design and experimental evaluation of network-assisted strategies for HTTP adaptive streaming. *In Proceedings of the 7th International Conference on Multimedia Systems*, page 3. ACM, 2016.
- [CDK+04] Miguel Castro, Peter Druschel, Anne-Marie Kermarrec, Animesh Nandi, Antony Rowstron et Atul Singh : SplitStream : High-Bandwidth Multicast in Cooperative Environments. *Operating Systems Review (ACM)*, 37, 08 2004.
- [CDZ97] K.L. Calvert, M.B. Doar et E.W. Zegura : Modeling Internet topology. *IEEE Communications Magazine*, 35(6):160–163, June 1997.
- [Cisa] Cisco : Cisco 8540 Wireless Controller. <https://www.cisco.com/c/en/us/products/wireless/8540-wireless-controller/>. (Accessed March 06, 2019).
- [Cisb] Cisco : Cisco VNI. <https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/air-highlights.html>. (Accessed March 06, 2019).

- [Cis07] Cisco : Configuring a Load-Balancing Scheme. [https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipswitch\\_cef/configuration/xs-3s/isw-cef-xe-3s-book/isw-cef-load-balancing.pdf](https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipswitch_cef/configuration/xs-3s/isw-cef-xe-3s-book/isw-cef-load-balancing.pdf), 2007.
- [CMK<sup>+</sup>13] Jiasi Chen, Rajesh Mahindra, Mohammad Amir Khojastepour, Sampath Rangarajan et Mung Chiang : A scheduling framework for adaptive video delivery over cellular networks. *In Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 389–400. ACM, 2013.
- [Coh03] Brahm Cohen : Incentives build robustness in bittorrent. *In Workshop on Economics of Peer-to-Peer Systems*, 2003.
- [CPZZ15] Federico Chiariotti, Chiara Pielli, Andrea Zanella et Michele Zorzi : QoE-aware Video Rate Adaptation algorithms in multi-user IEEE 802.11 wireless networks. *In the Proceedings of the IEEE International Conference on Communications (ICC 2015)*, 2015.
- [CRC<sup>+</sup>08] Meeyoung Cha, Pablo Rodriguez, Jon Crowcroft, Sue Moon et Xavier Amatriain : Watching Television over an IP Network. *In Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, IMC '08*, pages 71–84, New York, NY, USA, 2008. Association for Computing Machinery.
- [CRSZ02] Yang-Hua Chu, S.G. Rao, S. Seshan et Hui Zhang : A Case for End System Multicast. *IEEE Journal on Selected Areas in Communications*, 20(8):1456–1471, 2002.
- [CŞMC<sup>+</sup>21] Berat Can Şenel, Maxime Mouchet, Justin Cappos, Olivier Fourmaux, Timur Friedman et Rick McGeer : Shared internet-scale measurement platforms. *In NSF Workshop on Overcoming Measurement Barriers to Internet Research (WOMBIR 2021)*, Online, United States, janvier 2021.
- [DB95] L. Delgrossi et L. Berger : Internet Stream Protocol Version 2 (ST2) Protocol Specification - Version ST2+. RFC 1819, RFC Editor, August 1995.
- [Dee89] Steve Deering : Host extensions for IP multicasting. STD 5, RFC Editor, August 1989. <http://www.rfc-editor.org/rfc/rfc1112.txt>.
- [DF07] B. Donnet et T. Friedman : Internet Topology Discovery : a Survey. *Commun. Surveys Tuts.*, 9(4):56–69, octobre 2007.
- [DH17] S. Deering et R. Hinden : Internet Protocol, Version 6 (IPv6) Specification. STD 86, RFC Editor, July 2017.
- [DHS15] Kai Dong, Jun He et Wei Song : QoE-aware adaptive bitrate video streaming over mobile networks with caching proxy. *In Computing, Networking and Communications (ICNC), 2015 International Conference on*, pages 737–741. IEEE, 2015.
- [DLL<sup>+</sup>00] C. Diot, B.N. Levine, B. Lyles, H. Kassem et D. Balensiefen : Deployment issues for the ip multicast service and architecture. *Network, IEEE*, 14(1):78–88, Jan/Feb 2000.
- [DSA<sup>+</sup>11] Florin Dobrian, Vyas Sekar, Asad Awan, Ion Stoica, Dilip Joseph, Aditya Ganjam, Jibin Zhan et Hui Zhang : Understanding the impact of video quality on user engagement. *In ACM SIGCOMM CCR*, volume 41, pages 362–373. ACM, 2011.
- [FGB<sup>+</sup>15] Arsham Farshad, Panagiotis Georgopoulos, Matthew Broadbent, Mu Mu et Nicholas Race : Leveraging SDN to provide an in-network QoE measurement framework. *In Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*, pages 239–244. IEEE, 2015.



- [FN14] Hamid Farhadi et Akihiro Nakao : Rethinking flow classification in SDN. *In 2014 IEEE International Conference on Cloud Engineering*, pages 598–603. IEEE, 2014.
- [For79] James W. Forgie : ST - A Proposed Internet Stream Protocol. IEN 119, M.I.T. Lincoln Laboratory, September 1979.
- [Fou98] Olivier Fourmaux : *Les communications multipoints dans les réseaux haut débit multimédia : Le multicast en environnement IP sur ATM*. Theses, Université Pierre et Marie Curie - Paris VI, décembre 1998.
- [GCJM12] Monia Ghobadi, Yuchung Cheng, Ankur Jain et Matt Mathis : Trickle : Rate Limiting YouTube Video Streaming. *In Proceedings of the 2012 USENIX conference on Annual Technical Conference*, pages 17–17. USENIX Association, 2012.
- [GDS<sup>+</sup>03] Krishna P. Gummadi, Richard J. Dunn, Stefan Saroiu, Steven D. Gribble, Henry M. Levy et John Zahorjan : Measurement, Modeling, and Analysis of a Peer-to-Peer File-Sharing Workload. *SIGOPS Oper. Syst. Rev.*, 37(5):314–329, octobre 2003.
- [GEB<sup>+</sup>13] Panagiotis Georgopoulos, Yehia Elkhatib, Matthew Broadbent, Mu Mu et Nicholas Race : Towards network-wide QoE fairness using openflow-assisted adaptive video streaming. *In Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking*, pages 15–20. ACM, 2013.
- [Geo] GeolP : <http://www.maxmind.com/>. Accessed : 2021-08-06.
- [GFTM16] Marwan Ghanem, Olivier Fourmaux, Fabien Tarissan et Takumi Miyoshi : P2PTV Multi-channel Peers Analysis. *In The 18th Asia-Pacific Network Operations and Management Symposium*, Kanazawa, Japan, octobre 2016. IEEE.
- [GKM<sup>+</sup>16] Mojgan Ghasemi, Partha Kanuparth, Ahmed Mansy, Theophilus Benson et Jennifer Rexford : Performance Characterization of a Commercial Video Streaming Service. *In Proceedings of the 2016 ACM on IMC*, pages 499–511. ACM, 2016.
- [GS09] Mehmet H Gunes et Kamil Sarac : Resolving IP Aliases in Building Traceroute-Based Internet Maps. *IEEE/ACM Transactions on Networking*, 17(6):1738–1751, 2009.
- [GT00] Ramesh Govindan et Hongsuda Tangmunarunkit : Heuristics For Internet Map Discovery. *In Proc. INFOCOM '00*, 2000.
- [HG12] Rémi Houdaille et Stéphane Gouache : Shaping HTTP Adaptive Streams for a Better User Experience. *In Proceedings of the 3rd Multimedia Systems Conference*, pages 1–9. ACM, 2012.
- [HJM<sup>+</sup>14] Te-Yuan Huang, Ramesh Johari, Nick McKeown, Matthew Trunnell et Mark Watson : A Buffer-Based Approach to Rate Adaptation : Evidence from a Large Video Streaming Service. *In Proceedings of the 2014 ACM conference on SIGCOMM*, pages 187–198. ACM, 2014.
- [HLL<sup>+</sup>06] Xiaojun Hei, Chao Liang, Jian Liang, Yong Liu et Keith Ross : Insight into PPLive : A measurement study of a large-scale P2P IPTV system. *In WWW Workshop of IPTV services over World Wide Web*, May 2006.
- [Hop00] Christian E Hopps : Analysis of an equal-cost multi-path algorithm. RFC 2992 (Informational), novembre 2000.
- [HRBSD03] Martin Heusse, Frederic Rousseau, Gilles Berger-Sabbatel et Andrzej Duda : Performance anomaly of 802.11 b. *In INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 836–843. IEEE, 2003.

- [Hua19] Huawei : Per-flow and per-packet load balancing. <https://support.huawei.com/enterprise/en/doc/EDOC1100055041/ebc8ad42>, 2019. Accessed : 2021-08-06.
- [HVMF13] Hiep Hoang-Van, Takumi Miyoshi et Olivier Fourmaux : P2PTV Traffic Localization by Deep Packet Inspection. *In IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, pages 375–380, Honolulu, HI, United States, juillet 2013.
- [HVMF14] Hiep Hoang-Van, Takumi Miyoshi et Olivier Fourmaux : A Hierarchical P2P Traffic Localization Method with Bandwidth Limitation. *In IEEE International Conference on Communications, ICC 2014*, pages 3136–3141, Sidney, Australia, juin 2014. IEEE.
- [HVMMF13] Hiep Hoang-Van, Koki Mizutani, Takumi Miyoshi et Olivier Fourmaux : P2P Traffic Localization by Forcing Packet Loss. *International Journal of Networked and Distributed Computing*, 1(4):251–259, novembre 2013.
- [HVSMF14] Hiep Hoang-Van, Yuki Shinozaki, Takumi Miyoshi et Olivier Fourmaux : A Router-Aided Hierarchical P2P Traffic Localization Based on Variable Additional Delay Insertion. *IEICE Transactions on Communications*, E97-B(1):29–39, janvier 2014.
- [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.
- [JSM<sup>+</sup>16] Junchen Jiang, Vyas Sekar, Henry Milner, Davis Shepherd, Ion Stoica et Hui Zhang : CFA : a practical prediction system for video QoE optimization. *In Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation*, pages 137–150. USENIX Association, 2016.
- [JSZ12] Junchen Jiang, Vyas Sekar et Hui Zhang : Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive. *In Proceedings of the 8th international conference on Emerging networking experiments and technologies*, pages 97–108. ACM, 2012.
- [JTM15] Parikshit Juluri, Venkatesh Tamarapalli et Deep Medhi : SARA : Segment aware rate adaptation algorithm for dynamic adaptive streaming over HTTP. *In 2015 IEEE International Conference on Communication Workshop (ICCW)*, pages 1765–1770. IEEE, 2015.
- [Jun] Juniper : Understanding the algorithm used to load balance traffic on MX series routers. [https://www.juniper.net/documentation/en\\_US/junos/topics/concept/hash-computation-mpcs-understanding.html](https://www.juniper.net/documentation/en_US/junos/topics/concept/hash-computation-mpcs-understanding.html). Accessed : 2021-08-06.
- [KBJK<sup>+</sup>06] Ethan Katz-Bassett, John P. John, Arvind Krishnamurthy, David Wetherall, Thomas Anderson et Yatin Chawathe : Towards IP Geolocation Using Delay and Topology Measurements. *In Proc. ACM IMC '06*, 2006.
- [Key10] Ken Keys : Internet-Scale IP Alias Resolution Techniques. *ACM SIGCOMM Computer Communication Review*, 40(1):50–55, 2010.
- [KF13] Hyojoon Kim et Nick Feamster : Improving Network Management with Software Sefined Networking. *Communications Magazine, IEEE*, 51(2):114–119, 2013.
- [KH07] Sangmin Kim et Khaled Harfoush : Efficient Estimation of More Detailed Internet IP Maps. *In Proc. ICC '07*, 2007.
- [KHL13] Ken Keys, Young Hyun, Matthew Luckie et Kim Claffy : Internet-Scale IPv4 Alias Resolution with MIDAR. *IEEE/ACM Transactions on Networking*, 21(2):383–399, 2013.

- [KRP05] Thomas Karagiannis, Pablo Rodriguez et Konstantina Papagiannaki : Should Internet Service Providers Fear Peer-Assisted Content Distribution? *In Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*, IMC '05, page 6, USA, 2005. USENIX Association.
- [LB17] Matthew Luckie et Robert Beverly : The Impact of Router Outages on the AS-level Internet. *In Proc. ACM SIGCOMM '17*, 2017.
- [LBB<sup>+</sup>13] Matthew Luckie, Robert Beverly, William Brinkmeyer *et al.* : Speedtrap : internet-Scale IPv6 Alias Resolution. *In Proc. IMC '13*, 2013.
- [LCL11] Pochiang Lin, Wu-I Chou et Tsungnan Lin : Achieving airtime fairness of delay-sensitive applications in multirate IEEE 802.11 wireless LANs. *Communications Magazine, IEEE*, 49(9):169–175, 2011.
- [LDcM11] Matthew Luckie, Amogh Dhamdhere, kc claffy et David Murrell : Measured Impact of Crooked Traceroute. *SIGCOMM Comput. Commun. Rev.*, 41(1):14–21, janvier 2011.
- [LHH08] Matthew Luckie, Young Hyun et Bradley Huffaker : Traceroute probe method and forward IP path inference. *In Proc. ACM IMC '08*, 2008.
- [LKR06] Jian Liang, Rakesh Kumar et Keith W. Ross : The FastTrack overlay : A measurement study. *Computer Networks*, 50(6):842–858, 2006. Overlay Distribution Structures and their Applications.
- [LL16] Yan Liu et Jack YB Lee : A unified framework for automatic quality-of-experience optimization in mobile video streaming. *In IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9. IEEE, 2016.
- [Luc10] Matthew Luckie : Scamper : a Scalable and Extensible Packet Prober for Active Measurement of the Internet. *In Proc. IMC '10*, 2010.
- [MB94] M.R. Macedonia et D.P. Brutzman : MBone provides audio and video across the Internet. *Computer*, 27(4):30–36, April 1994.
- [Mer] Meraki : Meraki Cloud Controller. <https://goo.gl/R715X2>. (Accessed March 06, 2019).
- [MFA15] Ahmed Mansy, Marwan Fayed et Mostafa Ammar : Network-layer Fairness for Adaptive Video Streams. *In Proceedings of the IFIP Networking Conference*. IFIP, 2015.
- [MMF15] Koki Mizutani, Takumi Miyoshi et Olivier Fourmaux : Traffic analysis in concurrent multi-channel viewing on P2PTV. *Lecture notes in Electrical Engineering*, 339:85–92, 2015.
- [MPC<sup>+</sup>08] Rohan Murty, Jitendra Padhye, Ranveer Chandra, Alec Wolman et Brian Zill : Designing high performance enterprise Wi-Fi networks. *In Proceedings of the 5th USENIX Symposium on NSDI*, pages 73–88. USENIX Association, 2008.
- [MPP13] Pietro Marchetta, Valerio Persico et Antonio Pescapè : Pythia : Yet Another Active Probing Technique for Alias Resolution. *In Proc. CoNEXT '13*, 2013.
- [MSA<sup>+</sup>15] Mu Mu, Simpson Steven, Farshad Arsham, Ni Qiang et Race Nicholas : User-level Fairness Delivered : Network Resource Allocation for Adaptive Video Streaming. *In International Symposium on Quality of Service (IWQoS), 2015 IEEE/ACM Proceedings of*, pages 239–244. IEEE/ACM, 2015.

- [MTAA<sup>+</sup>16] Tarun Mangla, Nawanol Theera-Ampornpunt, Mostafa Ammar, Ellen Zegura et Saurabh Bagchi : Video through a crystal ball : effect of bandwidth prediction quality on adaptive streaming in mobile environments. *In Proceedings of the 8th International Workshop on Mobile Video*, page 1. ACM, 2016.
- [NKS16] Hyunwoo Nam, Kyung-Hwa Kim et Henning Schulzrinne : QoE matters more than QoS : Why people stop watching cat videos. *In Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on*, pages 1–9. IEEE, 2016.
- [NMF11] Yuko Nakamura, Takumi Miyoshi et Olivier Fourmaux : Traffic Analysis Considering Channel Zapping on P2P Video Delivery Service. *In IEICE*, numéro CQ2011-27, Hokkaido, Japan, juillet 2011. IEICE.
- [NS3] NS3 : NS3 Simulator Project. <https://www.nsnam.org>. (Accessed March 06, 2019).
- [ONF12] ONF : Software-Defined Networking : The New Norm for Networks. <https://opennetworking.org/sdn-resources/whitepapers/software-defined-networking-the-new-norm-for-networks/>, 2012.
- [PCB06] L. Plissonneau, J.-L. Costeux et P. Brown : Detailed analysis of eDonkey transfers on ADSL. *In 2nd Conference on Next Generation Internet Design and Engineering*, pages 8 pp.–262, April 2006.
- [PG98] Jean-Jacques Pansiot et Dominique Grad : On Routes and Multicast Trees in the Internet. *ACM SIGCOMM Computer Communication Review*, 28(1):41–50, 1998.
- [PM17] R. Pantos et W. May : HTTP Live Streaming. RFC 8216, RFC Editor, August 2017.
- [Pos81a] J. Postel : Internet Control Message Protocol. STD 5, RFC Editor, September 1981. <http://www.rfc-editor.org/rfc/rfc792.txt>.
- [Pos81b] Jon Postel : Internet Protocol. STD 5, RFC Editor, September 1981. <http://www.rfc-editor.org/rfc/rfc791.txt>.
- [PPL06] PPLive : <https://web.archive.org/web/20060104004558/http://pplive.com/>, 2006. Accessed : 2021-08-06.
- [PPS06] PPStream : <https://web.archive.org/web/20060220190144/http://ppstream.com/>, 2006. Accessed : 2021-08-06.
- [PRR<sup>+</sup>03] Saar Pilosof, Ramachandran Ramjee, Danny Raz, Yuval Shavitt et Prasun Sinha : Understanding TCP fairness over wireless LAN. *In INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 863–872. IEEE, 2003.
- [PWH<sup>+</sup>15] Stefano Petrangeli, Tim Wauters, Rafael Huysegems, Tom Bostoen et Filip De Turck : Network-based dynamic prioritization of HTTP adaptive streams to avoid video freezes. *In Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on*, pages 1242–1248. IEEE, 2015.
- [QLJ<sup>+</sup>13] Zafar Ayyub Qazi, Jeongkeun Lee, Tao Jin, Gowtham Bellala, Manfred Arndt et Guevara Noubir : Application-awareness in SDN. *In ACM SIGCOMM computer communication review*, volume 43, pages 487–488. ACM, 2013.
- [QZS16] Jason J Quinlan, Ahmed H Zahran et Cormac J Sreenan : Datasets for AVC (H. 264) and HEVC (H. 265) evaluation of dynamic adaptive streaming over HTTP (DASH). *In Proceedings of the 7th International Conference on Multimedia Systems*, page 51. ACM, 2016.

- [rip] RIPE Atlas. <https://atlas.ripe.net>.
- [RIP15] RIPE NCC Staff : RIPE Atlas. *Internet Protocol Journal*, 18(3), 2015.
- [ROF15] Vishwanath Ramamurthi, Ozgur Oyman et Jeffrey Foerster : Using link awareness for HTTP Adaptive Streaming over changing wireless conditions. *In Computing, Networking and Communications (ICNC), 2015 International Conference on*, pages 727–731. IEEE, 2015.
- [Roh20] Justin P. Rohrer : Operationalizing Yarrp : High-Speed Active Network Topology Mapping from AWS. Rapport technique, CAIDA, February 2020. [https://www.caida.org/workshops/kismet/2002/slides/kismet2002\\_jrohrer.pdf](https://www.caida.org/workshops/kismet/2002/slides/kismet2002_jrohrer.pdf).
- [RWM<sup>+</sup>11] Matthew Roughan, Walter Willinger, Olaf Maennel, Debbie Perouli et Randy Bush : 10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems. *IEEE Journal on Selected Areas in Communications*, 29(9):1810–1821, October 2011.
- [San20] Sandvine : The Global Internet Phenomena Report : COVID-19 Spotlight. <https://www.sandvine.com/covid-internet-spotlight-report>, May 2020.
- [SAR<sup>+</sup>09] Vivek Shrivastava, Nabeel Ahmed, Shravan Rayanchu, Suman Banerjee, Srinivasan Keshav, Konstantina Papagiannaki et Arunesh Mishra : CENTAUR : realizing the full potential of centralized wlans through a hybrid data path. *In Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 297–308. ACM, 2009.
- [SBS08] Rob Sherwood, Adam Bender et Neil Spring : Discarte : a Disjunctive Internet Cartographer. *ACM SIGCOMM Computer Communication Review*, 38(4):303–314, 2008.
- [SCFJ03] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson : RTP : A Transport Protocol for Real-Time Applications. STD 64, RFC Editor, July 2003. <http://www.rfc-editor.org/rfc/rfc3550.txt>.
- [ScH<sup>+</sup>11] James P.G. Sterbenz, Egemen K. Çetinkaya, Mahmood A. Hameed, Abdul Jabbar, Qian Shi et Justin P. Rohrer : Evaluation of Network Resilience, Survivability, and Disruption Tolerance : Analysis, Topology Generation, Simulation, and Experimentation. *Springer Telecommunication Systems*, 52(2):705–736, February 2011.
- [SF06] Thomas Silverston et Olivier Fourmaux : P2P IPTV Measurement : A Case Study of TVAnts. *In ACM CONEXT 2006*, pages 45 :1–45 :2, Lisboa, Portugal, décembre 2006. ACM.
- [SF07] Thomas Silverston et Olivier Fourmaux : Measuring P2P IPTV Systems. *In ACM NOSSDAV'07 - Network and Operating Systems Support for Digital Audio & Video*, Urbana-Champaign, IL, United States, juin 2007.
- [SFC08] Thomas Silverston, Olivier Fourmaux et Jon Crowcroft : Towards an Incentive Mechanism for Peer-to-Peer Multimedia Live Streaming Systems. *In IEEE P2P 2008 - The Eight International Conference on Peer-to-Peer Computing*, pages 125–128, Aachen, Germany, septembre 2008. IEEE.
- [SFS07] Thomas Silverston, Olivier Fourmaux et Kavé Salamatian : Characterization of TVAnts : Scaling Analysis of a P2P IPTV Traffic. *In MED-HOC-NET 2007*, Corfu, Greece, juin 2007.
- [SFS08] Thomas Silverston, Olivier Fourmaux et Kavé Salamatian : Caractérisation du trafic P2P IPTV : une analyse d'échelle. *In Colloque Francophone sur l'Ingénierie des Protocoles (CFIP 2008)*, Les Arcs, France, mars 2008.

- [SFS<sup>+</sup>09] Thomas Silverston, Olivier Fourmaux, Kavé Salamatian, Alberto Dainotti, Alessio Botta, Antonio Pescapè et Giorgio Ventre : Traffic Analysis of Peer-to-Peer IPTV Communities. *Computer Networks*, 53(4):470–484, mars 2009.
- [SFSC07] Thomas Silverston, Olivier Fourmaux, Kavé Salamatian et Kenjiro Cho : Measuring P2P IPTV Traffic on Both Sides of The World. In *ACM CoNEXT 2007 (CoNEXT)*, New York, NY, United States, décembre 2007. ACM.
- [SFSC09] Thomas Silverston, Olivier Fourmaux, Kavé Salamatian et Kenjiro Cho : Mesures d'applications P2P-TV de part et d'autre du monde. In *CFIP'2009*, Strasbourg, France, octobre 2009.
- [SFSC10] Thomas Silverston, Olivier Fourmaux, Kavé Salamatian et Kenjiro Cho : On Fairness and Locality in P2P-TV through Large-Scale Measurement Experiment. In *GLOBECOM 2010 - IEEE Global Telecommunications Conference*, pages 1–6, Miami, United States, décembre 2010.
- [SGMZ04] Kunwadee Sripanidkulchai, Aditya Ganjam, Bruce Maggs et Hui Zhang : The Feasibility of Supporting Large-Scale Live Streaming Applications with Dynamic Application End-Points. In *Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM '04*, pages 107–120, New York, NY, USA, 2004. Association for Computing Machinery.
- [Sil09] Thomas Silverston : *Peer-to-Peer Video Live Streaming : Measurement Experiments and Traffic Analysis*. Thèse de doctorat, 2009. Thèse de doctorat dirigée par Fdida, Serge Systèmes informatiques Paris 6 2009.
- [SJCA<sup>+</sup>11] Thomas Silverston, Lorand Jakab, Albert Cabellos-Aparicio, Olivier Fourmaux, Kavé Salamatian et Kenjiro Cho : Large-scale Measurement Experiments of P2P-TV Systems. Insights on Fairness and Locality. *Signal Processing : Image Communication*, 26(7):327–338, août 2011.
- [SKS09] Jan Seedorf, Sebastian Kiesel et Martin Stiemerling : Traffic localization for P2P-applications : The ALTO approach. In *2009 IEEE Ninth International Conference on Peer-to-Peer Computing*, pages 171–177, Sep. 2009.
- [SMC<sup>+</sup>21] Berat Can Şenel, Maxime Mouchet, Justin Cappos, Olivier Fourmaux, Timur Friedman et Rick McGeer : EdgeNet : A Multi-Tenant and Multi-Provider Edge Cloud. In *EdgeSys'21 - 4th International Workshop on Edge Systems, Analytics and Networking*, pages 49–54, Online, United Kingdom, avril 2021. ACM.
- [SMF12] Yuki Shinozaki, Takumi Miyoshi et Olivier Fourmaux : A P2P Traffic Localization Method with Additional Delay Insertion. In *4th International Conference on Intelligent Networking and Collaborative Systems, INCoS'2012*, pages 148–154, Bucarest, Romania, septembre 2012. IEEE.
- [SMG<sup>+</sup>07] A. Sentinelli, G. Marfia, M. Gerla, L. Kleinrock et S. Tewari : Will iptv ride the peer-to-peer stream? [peer-to-peer multimedia streaming]. *Communications Magazine, IEEE*, 45(6):86–92, June 2007.
- [SMW02] Neil Spring, Ratul Mahajan et David Wetherall : Measuring ISP Topologies with Rocketfuel. *ACM SIGCOMM Computer Communication Review*, 32(4):133–145, 2002.
- [SOP06] SOPCast : <https://web.archive.org/web/20060705070518/http://www.sopcast.com:80/>, 2006. Accessed : 2021-08-06.
- [SS15] Arkadeep Sen et Krishna M Sivalingam : An SDN framework for seamless mobility in enterprise WLANs. In *Personal, Indoor, and Mobile Radio Communications*

- (PIMRC), 2015 IEEE 26th Annual International Symposium on, pages 1985–1990. IEEE, 2015.
- [SZMC<sup>+</sup>15] Julius Schulz-Zander, Carlos Mayer, Bogdan Ciobotaru, Stefan Schmid et Anja Feldmann : OpenSDWN : Programmatic control over home and enterprise WiFi. *In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, page 16. ACM, 2015.
- [TCCP15] Daphne Tuncer, Marinos Charalambides, Stuart Clayman et George Pavlou : Adaptive resource management and control in software defined networks. *IEEE Transactions on Network and Service Management*, 12(1):18–33, 2015.
- [TLPR14] Truong Cong Thang, Hung T. Le, Anh T. Pham et Yong Man Ro : An Evaluation of Bitrate Adaptation Methods for HTTP Live Streaming. *IEEE Journal on Selected Areas in Communications*, 32(4):693–705, April 2014.
- [Tur95] Thierry Turletti : IVS (INRIA Videoconferencing System). <http://planete.inria.fr/ivs/>, 1995.
- [TVA06] TVAnts : <https://web.archive.org/web/20060101024130/http://www.tvants.com/>, 2006. Accessed : 2021-08-06.
- [TVU11] TVUPlayer : <https://web.archive.org/web/20110430055937/http://www.uusee.com/>, 2011. Accessed : 2021-08-06.
- [UUS07] UUSee : <https://web.archive.org/web/20070525165200/https://www.tvunetworks.com/>, 2007. Accessed : 2021-08-06.
- [VATF09] Darryl Veitch, Brice Augustin, Renata Teixeira et Timur Friedman : Failure Control in Multipath Route Tracing. *In IEEE INFOCOM 2009*, pages 1395–1403. IEEE, 2009.
- [Ver20] Kevin Vermeulen : *Improved Algorithms for Capturing Internet Maps*. Theses, Sorbonne Université, juin 2020.
- [Vet95] Ronald J. Vetter : ATM Concepts, Architectures, and Protocols. *Commun. ACM*, 38(2):30–ff., février 1995.
- [VLA<sup>+</sup>20] Kevin Vermeulen, Burim Ljuma, Vamsi ADDANKI, Matthieu Gouel, Olivier Fourmaux, Timur Friedman et Reza REJAIE : Alias Resolution Based on ICMP Rate Limiting. *In PAM 2020 - 21st International Conference on Passive and Active Network Measurement*, volume 12048 de *Lecture Notes in Computer Science*, pages 231–248, Eugene, United States, mars 2020. Springer.
- [VLF<sup>+</sup>19] Kevin Vermeulen, Burim Ljuma, Olivier Fourmaux, Timur Friedman et Rick McGeer : Internet measurements on EdgeNet. *In CNERT : Computer and Networking Experimental Research using Testbeds in conjunction with IEEE INFOCOM 2019*, Paris, France, avril 2019.
- [VPMD13] Yves Vanaubel, Jean-Jacques Pansiot, Pascal Mérindol et Benoit Donnet : Network Fingerprinting : TTL-Based Router Signatures. *In Proc. IMC '13*, 2013.
- [VR11] Silvio Valenti et Dario Rossi : Fine-grained behavioral classification in the core : the issue of flow sampling. *In 2011 7th International Wireless Communications and Mobile Computing Conference*, pages 1028–1032, July 2011.
- [VRB<sup>+</sup>20] Kevin Vermeulen, Justin P Rohrer, Robert Beverly, Olivier Fourmaux et Timur Friedman : Diamond-Miner : Comprehensive Discovery of the Internet’s Topology Diamonds. *In USENIX Symposium on Networked Systems Design and Implementation (NSDI) 2020*, pages 479–493, Santa Clara, CA, United States, février 2020. USENIX Association.

- [VSFF18] Kevin Vermeulen, Stephen D Strowes, Olivier Fourmaux et Timur Friedman : Multilevel MDA-Lite Paris Traceroute. *In IMC'18 - 2018 Internet Measurement Conference*, pages 29–42, Boston, United States, octobre 2018. ACM.
- [WNMF12] Manxue Wang, Yuko Nakamura, Takumi Miyoshi et Olivier Fourmaux : Network Impact of P2P-TV Zapping. *In WTC'2012*, Miyazaki, Japan, mars 2012.
- [Wro97] John Wroclawski : The Use of RSVP with IETF Integrated Services. RFC 2210, RFC Editor, September 1997. <http://www.rfc-editor.org/rfc/rfc2210.txt>.
- [XXCP15] Siyuan Xiang, Min Xing, Lin Cai et Jianping Pan : Dynamic rate adaptation for adaptive video streaming in wireless networks. *Signal Processing : Image Communication*, 39:305–315, 2015.
- [XYK+08] Haiyong Xie, Y. Richard Yang, Arvind Krishnamurthy, Yanbin Grace Liu et Abraham Silberschatz : P4p : Provider Portal for Applications. *In Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication*, SIGCOMM '08, pages 351–362, New York, NY, USA, 2008. Association for Computing Machinery.
- [YBC+14] Yiannis Yakoumis, Manu Bansal, Adam Covington, Johan van Reijndam, Sachin Katti et Nick McKeown : BeHop : a testbed for dense WiFi networks. *In Proceedings of the 9th ACM international workshop on Wireless network testbeds, experimental evaluation and characterization*, pages 1–8. ACM, 2014.
- [YJSS15] Xiaoqi Yin, Abhishek Jindal, Vyas Sekar et Bruno Sinopoli : A control-theoretic approach for dynamic adaptive video streaming over HTTP. *ACM SIGCOMM CCR*, 45(4):325–338, 2015.
- [ZEG+15] Xuan Kelvin Zou, Jeffrey Erman, Vijay Gopalakrishnan, Emir Halepovic, Rittwik Jana, Xin Jin, Jennifer Rexford et Rakesh K Sinha : Can accurate predictions improve video streaming in cellular networks? *In Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, pages 57–62. ACM, 2015.
- [ZLLY05] Xinyan Zhang, Jiangchuan Liu, Bo Li et Y.-S.P. Yum : CoolStreaming/DONet : a data-driven overlay network for peer-to-peer live media streaming. *In Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, volume 3, pages 2102–2111 vol. 3, March 2005.
- [ZPC+15] Yasir Zaki, Thomas Pötsch, Jay Chen, Lakshminarayanan Subramanian et Carmelita Görg : Adaptive congestion control for unpredictable cellular networks. *In ACM SIGCOMM CCR*, volume 45, pages 509–522. ACM, 2015.
- [ZQ16] Ahmed H Zahran et Jason J Quinlan : Impact of the LTE Scheduler on achieving Good QoE for DASH Video Streaming. *In Proceedings of the 22nd IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN 2016)*. IEEE, 2016.