



HAL
open science

Cartographie distribuée du coeur de l'Internet

Mickaël Hoerdt, Damien Magoni

► **To cite this version:**

Mickaël Hoerdt, Damien Magoni. Cartographie distribuée du coeur de l'Internet. *Annals of Telecommunications - annales des télécommunications*, 2005, 60 (5-6), pp.558-587. 10.1007/BF03219937. hal-01058189

HAL Id: hal-01058189

<https://hal.science/hal-01058189>

Submitted on 1 Aug 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.

Cartographie distribuée de la topologie du coeur de l'Internet

Mickaël Hoerd, Damien Magoni

*Université Louis Pasteur – LSIT
Boulevard Sébastien Brant
67400 Illkirch, France
{hoerd,magoni}@dpt-info.u-strasbg.fr*

RÉSUMÉ. L'étude de la topologie d'Internet permet de découvrir des propriétés pouvant être exploitées par des protocoles réseaux ou applicatifs afin d'améliorer leurs performances. La cartographie d'Internet est donc très profitable aux logiciels de simulation de protocoles réseaux. Malheureusement la cartographie d'Internet souffre de deux problèmes majeurs. D'abord elle prend du temps à être réalisée et ensuite elle souffre souvent d'incomplétude. Pour pallier ces deux handicaps, nous avons développé un logiciel de cartographie dont l'objectif est de cartographier le coeur d'Internet avec un maximum de précision en un minimum de temps. Dans cet article nous décrivons notre logiciel et évaluons ses performances par rapport à un logiciel de cartographie existant qui est en libre disposition pour la recherche. Nous évaluons la complétude des cartes de routeurs en effectuant des recouvrements avec une topologie de systèmes autonomes. Nous évaluons aussi son utilisation dans le cadre d'IPv6 ainsi que l'évolution de ses performances en IPv4 en fonction des paramètres d'entrées tels que le nombre de traceurs utilisés. Enfin nous proposons une étude détaillée des délais des chemins et des liens dans l'Internet.

ABSTRACT. The study of the Internet topology allows us to discover properties that can be beneficial to the performances of network protocols and applications. Thus, Internet mapping is very useful to simulation software focusing on network protocols. Unfortunately Internet mapping has two major drawbacks. First it takes a lot of time to be carried out and second it is often incomplete. To solve these two problems, we have developed a fast mapping software whose aim is to map the heart of the Internet as fast as possible with the highest attainable precision. In this paper, we describe our software and evaluate its performances compared to an existing and freely available mapping software. We assess the completeness of the router level maps by creating overlays with an autonomous system topology. We also assess its use in IPv6 and the evolution of its performances in IPv4 when increasing the size of the input parameter set. We conclude with a detailed study on the delay of the paths and links in the Internet.

MOTS-CLÉS : carte, délai, Internet, routeur, système autonome, topologie.

KEYWORDS: autonomous system, delay, Internet, map, router, topology.

1. Introduction

Nous décrivons dans cet article notre logiciel de cartographie rapide de l'Internet ainsi que notre méthode de construction de recouvrements routeurs-systèmes autonomes ou AS (Autonomous Systems en anglais) afin d'évaluer la complétude des cartes de routeurs. La cartographie d'Internet est utile pour la simulation de protocoles destinés à être déployés dans l'Internet. Elle est en particulier très utile à la simulation des protocoles situés au niveau de la couche réseau (couche 3 du modèle OSI) tels que les protocoles de routage, les protocoles multipoints ou encore les protocoles de gestion de la qualité de service. L'article est organisé en huit sections. La section 2 présente les travaux similaires au nôtre ainsi que le logiciel que nous utilisons pour pouvoir effectuer une comparaison avec notre logiciel. La section 3 présente le fonctionnement de notre logiciel de cartographie. La section 4 présente les cartes collectées par les deux logiciels ainsi que la carte des systèmes autonomes utilisée pour permettre de créer les recouvrements. La section 5 expose les limitations liées à l'utilisation des cartes. La section 6 décrit la méthodologie de construction d'un recouvrement. La section 7 présente des résultats d'analyse concernant la précision des cartes de routeurs et donc de l'efficacité des logiciels de cartographie. La section 8 présente les techniques utilisées pour cartographier le réseau IPv6 ainsi que des premiers résultats comparatifs entre IPv6 et IPv4. Enfin la section 9 présente les résultats d'une deuxième campagne IPv4 à très grande échelle. Elle décrit les limites de l'approche quantitative de notre outil et présente une étude détaillée sur les longueurs et les délais des chemins ainsi que sur les délais des liens de cette nouvelle carte.

2. Travaux similaires

La cartographie de l'Internet fait l'objet de recherches depuis plusieurs années. En 1995, Pansiot *et al.* ont créé une carte de routeurs avec des *traceroute* provenant d'une machine unique vers environ 5000 destinations [PAN 98]. Ils ont utilisé le routage par source (*source routing*) pour découvrir les liens transverses. En 1999, Burch *et al.* ont utilisé des tables provenant du protocole de routage inter-domaine BGP (Border Gateway Protocol) pour trouver des préfixes destinations [BUR 99]. Ils ont aussi effectué des *traceroute* à partir d'une seule machine mais ils ont utilisé des tunnels vers d'autres machines obtenant ainsi un effet similaire au *source routing*. Toujours en 1999, Govindan *et al.* ont créé un logiciel de cartographie nommé Mercator qui utilise une méthode aléatoire de sélection d'adresses destinations [GOV 00]. C'est ce logiciel que nous utilisons pour le comparer à *nec* car c'est le seul logiciel de cartographie disponible gratuitement pour la recherche. Il fonctionne à partir d'une machine unique en envoyant des paquets UDP (User Datagram Protocol, envoi en mode non-connecté) avec un TTL (Time To Live, champ d'un paquet IP indiquant combien ce dernier a traversé de routeurs) croissant (donc similaire à l'utilisation d'une commande *traceroute*). Il utilise aussi le *source routing* pour découvrir les liens transverses. En 2002, Spring *et al.* ont créé un logiciel nommé Rocketfuel [SPR 02]. Ce logiciel fonctionne de manière similaire à *nec* (i.e. il interroge des serveurs *traceroute*) mais il ne permet

de cartographier qu'un seul ISP (Internet Service Provider, fournisseur d'accès) à la fois. De plus il n'est pas disponible, par contre le logiciel de résolution des alias IP utilisé par Rocketfuel et nommé Ally est disponible et nous l'utilisons avec *nec*.

Concernant notre méthode de création de recouvrement, celle-ci est différente et innovante par rapport aux méthodes utilisées dans [TAN 01b] et dans [CHA 01] car nous faisons correspondre les routeurs découverts par Mercator ou *nec* aux systèmes autonomes trouvés dans la table BGP de *route-views* grâce à l'utilisation des interfaces IP et des préfixes BGP. Par conséquent nous n'avons pas besoin de générer le graphe des AS par un algorithme de regroupement tel que celui présenté dans [TAN 01b] et nous évitons les erreurs potentielles introduites par les cas où beaucoup de groupes disjoints de routeurs appartenant au même AS doivent être réassignés. De plus notre méthode de création, contrairement aux méthodes précédentes, présente l'avantage de pouvoir évaluer le taux de recouvrement (pour les AS et les connexions BGP) de la carte de niveau AS par celle de niveau routeur. Cela nous permet d'estimer la qualité de la complétude de la carte de niveau routeur. Nous définissons la complétude comme étant un rapport variant entre 0 et 1 (ou exprimé en pourcentage) et indiquant quelle est la proportion des AS (resp. des connexions BGP) contenant au moins 1 routeur (resp. 1 lien IP) parmi tous ceux de la carte des AS.

3. network cartographer

Notre logiciel de cartographie de l'Internet se nomme *network cartographer* ou *nec* en abrégé. *nec* se lance depuis un point donné de l'Internet (typiquement sur la machine d'un utilisateur) et interroge simultanément un ensemble de serveurs *traceroute* vers un ensemble de destinations stockées localement. Le cardinal de ces ensembles pouvant potentiellement être très élevé, un mécanisme de contrôle de ressources au niveau local et distant est mis en oeuvre. A un instant donné, le nombre de requêtes en cours est limité. Typiquement, sur un ensemble de 76 serveurs *traceroute* et environ 6000 adresses destinations nous avons décidé de limiter le nombre local de requêtes simultanées à 20. Le nombre de requêtes distantes simultanées provenant de la machine sur laquelle s'exécute *nec* est limité à 1 pour ne pas surcharger un serveur plus que ne le ferait une requête manuelle faite via un navigateur web. Dans la suite du texte nous appellerons la limite locale x .

De manière plus détaillée, Le fonctionnement est le suivant : à l'initialisation, on ouvre x connexions HTTP vers x serveurs *traceroute* distincts auxquels on soumet une requête vers une même destination y . On attend ensuite les résultats, sachant qu'on ne peut pas prévoir l'ordre dans lequel ils arriveront et qu'ils peuvent être incomplets en raison des propriétés inhérentes de l'Internet (i.e. filtrage des paquets ICMP sur certains routeurs).

A réception d'une ou plusieurs réponses, on les enregistre pour les traiter *offline*. Si une réponse est complète, on ferme la connexion pour en ouvrir une autre et envoyer une nouvelle requête soit vers une adresse destination égale à y à un serveur non

encore sollicité pour cette adresse, soit vers une adresse destination différente de y à un serveur ayant déjà traité y . Si la réponse est incomplète, on récupère les données déjà envoyées et on se remet en attente en se basant sur les mécanismes qu'offre TCP pour gérer les erreurs. On ferme la connexion si une requête *traceroute* est filtrée (i.e. si on reçoit une succession de symboles *). Avant toute ouverture de nouvelle connexion, on vérifie nos deux critères d'économie de ressources : qu'une requête n'est pas déjà en cours sur le serveur *traceroute* qu'on aimerait interroger et que le nombre de requêtes en cours est inférieur ou égal à x . *nec* stocke des informations d'états afin de pouvoir reprendre une cartographie interrompue sans avoir à recommencer depuis le début.

Le choix des adresses de destination y (i.e. cibles des *traceroute*) est basé sur le système autonome auquel y appartient. Les adresses destinations appartiennent toutes à un AS distinct afin d'obtenir une distribution topologique optimale des cibles et par conséquent une carte la plus précise possible en utilisant le moins de requêtes possible.

nec est écrit en C et en C++. Son code source est disponible gratuitement pour toute utilisation non commerciale. Il peut être téléchargé à l'adresse [MAG a]. Cependant il n'est pas encore terminé et nous souhaitons y implémenter de nombreuses fonctionnalités supplémentaires. Il est mis à disposition de la communauté de recherche en réseaux essentiellement à des fins d'évaluation.

4. Collecte des cartes d'Internet

Notre objectif est de cartographier rapidement le coeur de l'Internet. Les campagnes de cartographie des travaux précédents [PAN 98, BUR 99, GOV 00] s'effectuent en général sur plusieurs semaines voire plusieurs mois. Ici nous effectuons une campagne de mesures dont la durée est de 10 jours. Bien que cela paraisse déjà important, il ne faut pas oublier qu'une campagne plus rapide pourrait porter atteinte au service fourni par les serveurs *traceroute* (dans le cas de *nec*) et/ou pourrait être interprétée comme une attaque de déni de service distribuée vis-à-vis des adresses de destination (Mercator et *nec*) ou encore comme une tentative d'espionnage de réseau (Mercator). De plus nous parlons d'une cartographie du *coeur* de l'Internet parce que nous sommes limités aux adresses IP publiques (on ne peut pas cartographier les adressages privés accessibles à travers des boîtes NAT (Network Address Translation)) et nous sommes dépendant du traitement des paquets ICMP par les routeurs. Certains, suite à une sollicitation, ne répondent pas par renvoi de paquets ICMP, d'autres filtrent carrément tout paquet ICMP. Ce dernier comportement est très fréquent dans les réseaux des fournisseurs d'accès Internet grand public, mais très rare chez les gros opérateurs. Il en résulte que l'on cartographie plutôt le coeur que les extrémités du réseau (i.e. les *traceroute* n'atteignent pas toujours leurs cibles). Nous avons donc fait fonctionner Mercator et *nec* pendant 10 jours et avons comparé les cartes de routeurs ainsi obtenues.

Pour créer les recouvrements nécessaires à l'analyse de la complétude, nous avons besoin des cartes des routeurs de L'Internet et d'une carte des systèmes autonomes.

Tableau 1. *Cartes Internet utilisées pour l'analyse*

Logiciel ou origine	Nb d'interfaces ou de préfixes	Nb de routeurs ou d'AS	Nb de liens	Période
Mercator	58119	55425	65237	7-17/4/2003
<i>nec</i>	50123	47055	119909	19-29/4/2003
<i>route-views</i>	130815	15129	31378	1/4/2003

Nous avons collecté les 2 cartes de niveau routeur à partir d'une machine localisée dans notre laboratoire (le LSIIT, situé à Illkirch, France) en utilisant le logiciel Mercator développé par Govindan *et al.* et décrit dans [GOV 00] d'une part et notre logiciel *nec* d'autre part. Ces cartes contiennent les adresses IP des interfaces des routeurs. La résolution des alias IP (i.e. l'identification des interfaces appartenant à un même routeur) est une étape importante de la cartographie. Dans Mercator, celle-ci est réalisée parallèlement à la cartographie par une technique initialement proposée par Pansiot *et al.* dans [PAN 98]. Le logiciel envoie un paquet UDP vers un port fermé à chacune des 2 interfaces suspectées d'appartenir au même routeur. Si les adresses sources des paquets ICMP de réponse sont identiques, alors il y a une forte chance pour que ces 2 interfaces appartiennent au même routeur. Dans *nec*, la résolution des alias s'effectue un peu en décalé par un logiciel nommé Ally. Une liste des interfaces susceptibles d'être des alias est créée par *nec*. Sont suspectes, toutes les adresses IP ayant un suffixe DNS identique (lorsqu'il existe), se trouvant dans le même AS et découvertes par des traces différentes. Ally est décrit dans [SPR 02] et fonctionne de la façon suivante. Il envoie un paquet UDP vers un port fermé à chacune des 2 interfaces suspectées d'appartenir au même routeur. Si les valeurs du champ identificateur IP (IP identifier en anglais) des paquets ICMP de réponse sont très proches, alors il envoie un troisième paquet à la première interface. Si la troisième valeur est proche des deux premières et que les trois respectent un ordre croissant alors il y a une forte chance pour que ces 2 interfaces appartiennent au même routeur (i.e. l'identificateur IP est produit à partir du même compteur donc très vraisemblablement du même routeur).

Nous avons aussi récupéré une carte de niveau AS produite à partir d'une image de la table de routage du routeur BGP observateur *route-views* [Uni] au début d'avril 2003. Le tableau 1 contient des informations générales sur ces cartes. On remarque que Mercator a détecté 4,6% d'alias et Ally 6,1%. Sans préjuger de l'efficacité de chacune des deux méthodes de résolution d'alias, on peut noter que le taux d'alias détectés se situe autour de 5% ce qui est presque le tiers de la valeur observée dans [BAR 01].

5. Limitations inhérentes aux cartes

Avant de présenter la construction et l'analyse des recouvrements routeurs-AS, deux remarques doivent être formulées concernant les limitations majeures de notre étude :

– Les cartes Internet que nous utilisons sont incomplètes. La carte de niveau routeur créée par Mercator est incomplète pour les raisons présentées dans [GOV 00] et dans la section 7. Celle créée par *nec* est limitée par les défauts inhérents à la commande *traceroute*. La carte de niveau AS provenant des données du routeur BGP *route-views* est aussi incomplète particulièrement en ce qui concerne le nombre de connexions BGP. En effet, Chang *et al.* ont montré dans [CHA 02] que l'utilisation de sources additionnelles d'information de routage de niveau AS permet d'augmenter le nombre de connexions BGP détectées jusqu'à 47,7%. Cela a un impact direct sur la connectivité du réseau d'AS. Notre but n'est pas d'obtenir une précision absolue dans la cartographie mais plutôt et surtout une précision relative la plus élevée possible par rapport aux outils et aux données dont nous disposons pour la réaliser.

– Les cartes de niveau routeur produites par Mercator ou *nec* (ou n'importe quel autre outil de mesure topologique de niveau IP) ne doivent pas être vues comme contenant uniquement des routeurs réels et des liens point-à-point physiques mais simplement comme des cartes contenant des noeuds IP et des liens IP. D'ailleurs beaucoup de ces noeuds et liens IP sont purement virtuels et donc n'existent pas physiquement. Cela s'explique par le fait que, particulièrement dans les coeurs de réseaux, les technologies de commutation de circuits (virtuels) telles que Frame Relay, ATM ou MPLS, peuvent non seulement masquer leur propre topologie à la couche IP mais peuvent aussi émuler des routeurs ou des liens IP selon une topologie complètement différente de la leur. Bien que nous utilisions le terme "routeur" dans la suite de cet article pour des raisons de simplification, cette caractéristique ne doit pas être oubliée.

6. Construction des recouvrements

Nous proposons dans cette section une méthode de construction d'un recouvrement (*overlay* en anglais) routeurs–systèmes autonomes. Nous avons implémenté notre méthode dans notre logiciel *nem* [MAG b] dont le code est disponible sur le web. Nous espérons que de tels recouvrements seront utilisés par la communauté de recherche concernée par les protocoles de l'Internet. En effet, ces informations peuvent être utilisées pour optimiser le placement intra ou inter domaine de serveurs mandataires (*proxy* en anglais), de membres d'arbres multipoints applicatifs, etc. Cependant dans cet article, nous souhaitons utiliser les recouvrements pour évaluer la complétude des cartes de routeurs.

Nous construisons un recouvrement de manière à lier la topologie des routeurs à celle des AS. Nous utilisons le contenu d'une table de routage BGP créée le 1 avril 2003 par *route-views* pour fabriquer le recouvrement. Cette table contient 15129 AS. Pour la construction du recouvrement, nous associons chaque préfixe réseau trouvé dans la table à l'AS qui le déclare (i.e. l'AS situé le plus à droite du chemin d'AS). Cet AS n'est pas nécessairement l'AS d'origine du préfixe car ce dernier peut être masqué par l'aggrégation des chemins d'AS [REK 95]. Des erreurs peuvent donc être introduites à ce niveau. Lorsqu'un préfixe peut être associé à plus d'un AS (suite à une erreur de protocole ou de base de données), nous conservons le premier AS ayant

le drapeau "i" (i.e. interne) positionné si il existe, sinon nous conservons le premier AS trouvé (quelques cas dans notre table, aussi sources d'erreurs). La table contient 130815 préfixes réseaux (ce qui est consistant avec les résultats trouvés par Bu *et al.* dans [BU 02a]).

Puis nous utilisons nos informations sur la topologie IP collectées en utilisant le logiciel Mercator ou notre logiciel *nec* pour construire une carte de niveau routeur de l'Internet. Pour chaque interface, nous recherchons le plus long préfixe réseau correspondant et nous associons l'AS source (ou celui qui le déclare) possédant ce préfixe réseau à l'interface. Contrairement à la méthode utilisée dans [TAN 01b], nous n'avons pas utilisé les IRR (Internet Routing Registries, enregistrement des informations de routage) comme sources additionnelles d'information car elles ne sont pas assez précises pour notre utilisation. D'ailleurs Chen *et al.* ont montré dans [CHE 02] que presque 62% des enregistrements de la base de données RIPE (Réseaux IP Européens, base de données contenant en particulier des informations sur les plans d'adressage des réseaux européens) sont égaux à néant (*void* en anglais) ou obsolètes malgré le fait que RIPE soit activement maintenue à jour. Parmi les interfaces non résolues, beaucoup d'entre elles appartiennent à des AS (comme peuvent le montrer quelques requêtes aux IRR) mais certaines d'entre elles telles que les 188.1.x.x du réseau de recherche allemand (nommé DFN) sont configurées de sorte à n'appartenir à aucun AS. Nous marquons toutes les interfaces non résolues comme appartenant à l'AS de numéro 0. Nous définissons l'AS de numéro 0 comme suit : "une adresse IP associée à l'AS de numéro 0 n'appartient à aucun AS". En dépit du nombre important d'adresses privées trouvées dans [TAN 01b], nous n'avons pas trouvé d'adresses privées dans nos données de niveau IP produites par Mercator. Ces adresses sont très probablement filtrées sur le chemin de retour à notre machine.

Finalement pour chaque routeur, nous définissons son adresse IP et recherchons son numéro d'AS en utilisant les interfaces associées à chaque routeur par Mercator ou *nec*. Si le routeur n'a qu'une interface, le processus est simple : nous assignons l'adresse IP et le numéro d'AS de l'interface au routeur. Si le routeur possède plusieurs interfaces, nous recherchons l'interface ayant la plus petite adresse IP et un numéro d'AS différent de 0 et nous assignons cette adresse IP et son numéro d'AS à ce routeur (cette méthode peut produire des erreurs). A la fin du processus, chaque routeur possède une adresse IP par défaut ainsi qu'un numéro d'AS (qui peut être égal à 0). Nous lions alors les routeurs à leurs AS respectifs dans notre logiciel *nem* afin de créer la structure de données du recouvrement. Nous utilisons le recouvrement pour déterminer quels routeurs pourraient être des routeurs BGP potentiels. Pour chaque routeur, nous regardons ses routeurs voisins et si l'un ou plusieurs d'entre eux se trouve dans un AS ayant un numéro d'AS non nul et différent de celui du routeur initial alors nous marquons ce dernier comme étant un routeur BGP.

7. Analyse de la qualité des cartes

Dans cette section, nous étudions les propriétés des cartes collectées par Mercator et *nec* afin d'évaluer leurs qualités.

7.1. Propriétés générales

Le tableau 2 contient les valeurs des principales propriétés des cartes qui, nous le rappelons, ont été collectées en 10 jours. Premier point, *nec* a découvert environ 15% moins de routeurs que Mercator. Une augmentation de la fréquence des requêtes et/ou du nombre de serveurs *traceroute* nous permettrait sans doute de combler ce retard. Après construction des recouvrements, nous trouvons que 12% des routeurs présents dans la topologie IP de Mercator peuvent être définis comme des routeurs BGP selon notre méthode de résolution présentée à la section 6, alors que ce taux est de presque 60% pour la topologie recueillie par *nec*. Nous ne possédons pas actuellement d'informations permettant de vérifier ce chiffre par rapport à sa valeur réelle dans l'Internet, mais cela signifie que notre méthode de distribution des cibles (i.e. adresses de destination des *traceroute*) qui consistent à en choisir une par AS, fait découvrir beaucoup de liens IP inter-domaines. Deuxième point et certainement le plus important de tous, *nec* a découvert presque 84% de liens de plus que Mercator ! Ce résultat recherché par notre logiciel est très positif. Il illustre l'avantage de l'utilisation de points de mesures multiples et distribués par rapport aux techniques de *source-routing* qui sont très insuffisantes pour découvrir des liens transverses en peu de temps. Il s'ensuit un degré moyen de la carte *nec* plus de 2 fois supérieur à celui de Mercator (donc une beaucoup plus forte densité de liens redondants). Le pourcentage de la partie maillée (i.e. le graphe amputé de tous ses arbres) des cartes est éloquent à cet égard. L'Internet est beaucoup plus maillé que ne le faisait apparaître les travaux précédents (85% des routeurs de la carte *nec* sont dans le maillage). En tout cas la valeur du pourcentage de la partie maillée ne doit pas être utilisé comme indicateur (i.e. invariant d'une propriété du graphe). Toujours grâce aux recouvrements avec la carte d'AS, nous pouvons calculer la complétude des cartes de routeurs. La complétude des AS donne le pourcentage des AS qui contiennent au moins 1 routeur (idéalement, il devrait être de 100%). La complétude des connexions BGP donne le pourcentage des connexions BGP qui contiennent au moins 1 lien IP (il devrait être de 100% comme pour les AS). Ces indicateurs sont relativement grossiers mais suffisants pour montrer les limites des cartes de routeurs. On peut voir dans ces valeurs le fort biais de Mercator qui ne remplit que 12% des AS et 6% des connexions. *nec* par contre atteint des valeurs très supérieures avec des taux de remplissage respectifs de 35% et 28%. Bien que cela soit loin d'être exhaustif, le progrès apporté par *nec* par rapport aux techniques de cartographie antérieures est indéniable.

Tableau 2. *Propriétés des cartes collectées*

Propriété	carte Mercator	Carte <i>nec</i>	Ecart
Routeurs	55425	47055	-15.1%
Routeurs BGP	12.0%	59.8%	+398.3%
Liens IP	65237	119909	+83.8%
Degré moyen	2.35	5.09	+116.6%
Partie maillée	24.8%	85.7%	+245.5%
Complétude AS	11.7%	35.0 %	+199.1%
Complétude connexion BGP	6.1%	27.8%	+355.7%

Tableau 3. *Valeurs des principaux indicateurs*

Indicateur	carte Mercator	Carte <i>nec</i>
CC du degré	0.98	0.96
Exposant du degré	-2.80	-2.36
CC du rang	0.96	0.95
Exposant du rang	-0.65	-0.73
CC de la taille des arbres	0.98	0.97
Exposant de la taille des arbres	-2.12	-3.92
CC du rang des arbres	0.98	0.96
Exposant du rang des arbres	-0.83	-0.35
CC du nb de plus courts chemins	0.67	0.95
Exposant du nb de plus courts chemins	-1.38	-1.51
CC du rang des paires	0.99	0.99
Exposant du rang des paires	-0.58	-0.58
Distance moyenne	16.3	9.0
Excentricité	30.9	18.9
Rayon	23	15
Diamètre	44	28

7.2. Lois puissances et distances

Les indicateurs topologiques sont des outils intéressants pour évaluer la similitude d'un graphe avec la topologie de l'Internet. Les indicateurs sont définis dans nos travaux antérieurs [MAG 02a]. Le tableau 3 contient les valeurs des principaux indicateurs de nos 2 cartes de routeurs. Ils sont proches en général (des variations plus fortes sur les exposants sont usuelles et difficiles à interpréter), excepté ceux liés au nombre de plus courts chemins et ceux concernant les distances.

Le coefficient de corrélation (CC) du nombre de plus courts chemins (illustrant la présence de la loi puissance n°5 décrite dans [MAG 01]) n'est pas vérifié par la carte de Mercator (i.e. il est inférieur à la valeur seuil de 0.95). Cette propriété est liée à la

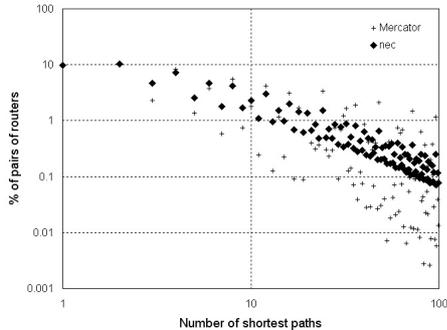


Figure 1. *Distr. du nombre de plus courts chemins.*

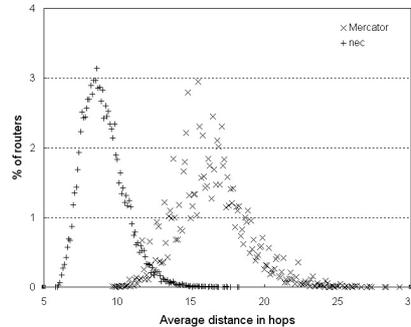


Figure 2. *Distr. de la distance moyenne des routeurs entre eux.*

quantité de liens redondants et nous pouvons illustrer par ce défaut le manque de précision de Mercator. La carte de *nec* par contre vérifie cette propriété cruciale grâce au nombre important de liens découverts. La figure 1 illustre la distribution du pourcentage de paires de routeurs ayant de 1 à 100 plus courts chemins. Les points doivent être alignés. On remarque la forte dissémination des points de Mercator alors que ceux de *nec* sont nettement plus rapprochés. Concernant les distances (exprimées en nombre de sauts, c.a.d. de liens traversés), la carte de Mercator a des propriétés très différentes des grandes cartes de routeurs alors que celle de *nec* possède pratiquement les mêmes valeurs que celles mesurées dans des cartes de plus de 200k routeurs [MAG 02b]. Le manque de liens de la carte de Mercator entraîne un étirement des distances entre les routeurs ce qui donne des valeurs augmentées de plus de 50% par rapport aux valeurs usuelles pour tous les indicateurs de distance. La figure 2 illustre la distribution du pourcentage de routeurs ayant une distance moyenne donnée par rapport aux autres (arrondie au dixième). Les points de *nec* forment une distribution normale très nette et bien centrée autour de 9 sauts contrairement à ceux de Mercator qui sont éparpillés autour de 17 sauts.

7.3. Présence de distributions à queue lourde

Pour vérifier la précision de nos recouvrements, nous étudions la distribution de la taille des AS. La taille d'un AS est égal au nombre de routeurs contenus dans cet AS. Nous avons montré précédemment que le taux de complétude est de 12% pour Mercator et de 35% pour *nec*. Cela montre que nos cartes de niveau routeur sont encore loin de la complétude totale. Nous cherchons les distributions à queue lourde trouvées dans les recouvrements d'Internet par Tangmunarunkit *et al.* dans [TAN 01a]. Une méthode éprouvée et utilisée dans [TAN 01a] et [BU 02b] pour réaliser cela consiste à tracer la fonction de distribution cumulative complémentaire de la taille des AS (CCDF en anglais) sur une double échelle logarithmique. Les figures 3 et 4 présentent

les CCDF des distributions de la taille des AS pour les cartes de Mercator et de *nec* et celles-ci présentent clairement le caractère d'une distribution à queue lourde. On peut noter que la CCDF de Mercator est légèrement bombée alors que celle de *nec* est parfaitement droite ce qui dénote une parfaite corrélation. Par contre la queue de la distribution de *nec* est brisée à partir d'une taille de 1000. Cet effet de bord est probablement dû à un manque de données causé par la courte période de mesure. Ce phénomène apparait aussi pour Mercator mais il est un peu masqué par la courbure du tracé.

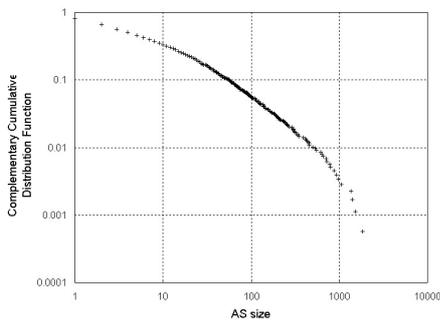


Figure 3. *Distr. cum. comp. de la taille des AS vus par Mercator.*

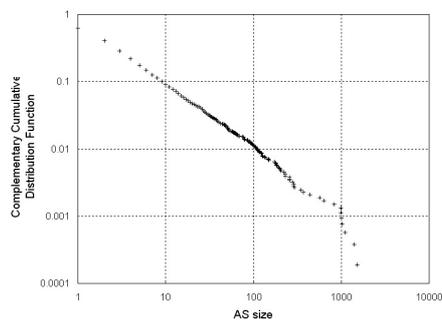


Figure 4. *Distr. cum. comp. de la taille des AS vus par nec.*

Nous tracons aussi sur les figures 5 et 6 la corrélation entre la taille et le degré des 100 plus gros AS. La corrélation pour ces AS est forte et très similaire à celle trouvée dans [TAN 01a] surtout pour la carte de *nec* où la linéarité est meilleure que pour celle de Mercator.

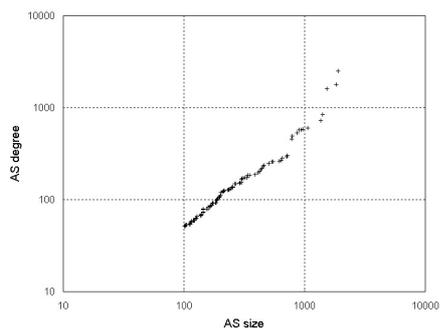


Figure 5. *Corrélation entre les tailles des AS et leurs degrés vus par Mercator.*

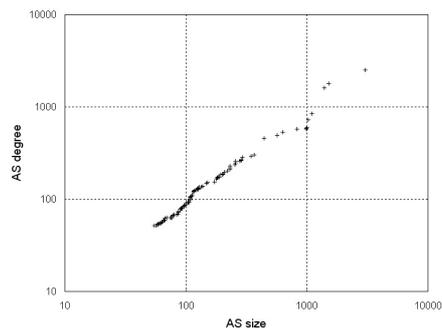


Figure 6. *Corrélation entre les tailles des AS et leurs degrés vus par nec.*

Les figures 7 et 8 présentent sur une double échelle logarithmique les CCDF des distributions du nombre de routeurs BGP par AS. Les formes sont très similaires à celles des tailles des AS et appellent les mêmes commentaires. Le nombre beaucoup plus important de routeurs BGP dans la carte de *nec* entraîne une linéarité presque parfaite, excepté pour des abscisses très élevées. Il est clair au vu des tracés que ces distributions sont fortement corrélées à celles de la taille des AS.

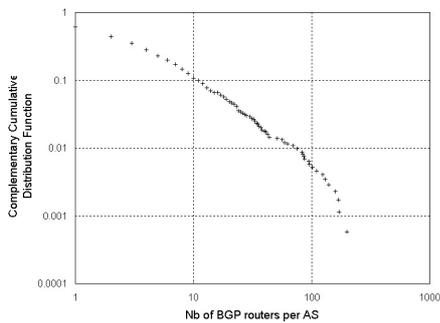


Figure 7. *Distr. cum. comp. du nb de routeurs BGP par AS vus par Mercator.*

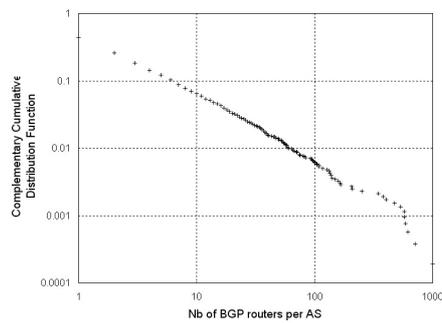


Figure 8. *Distr. cum. comp. du nb de routeurs BGP par AS vus par nec.*

Grâce à notre technique de création de recouvrement, nous pouvons étudier les tailles des connexions BGP. La taille d'une connexion BGP est égale au nombre de liens IP qui vont de n'importe quel routeur d'un AS donné vers n'importe quel routeur d'un AS donné différent du précédent. De façon similaire aux AS, seulement une partie des connexions contiennent un ou plusieurs liens IP. Ceci nous donne un ratio du nombre de connexions remplies (i.e. ayant une taille ≥ 1) vs le nombre total de connexions de 6% pour Mercator et 28% pour *nec* ce qui est plus faible que les taux de remplissage des AS. Une fois de plus cela tend à montrer que la carte de Mercator manque un nombre significatif de liens IP. Les figures 9 et 10 présentent sur une double échelle logarithmique les CCDF des distributions de la taille des connexions BGP et celles-ci révèlent des queues lourdes. Nous sommes, à notre connaissance, les premiers à présenter ce type de résultats (intuitivement prévisible).

Nous pouvons en conclure que bien que nos recouvrements routeur-AS soient incomplets, ils contiennent beaucoup de distributions à queue lourde typiquement trouvées au niveau macroscopique de la topologie de l'Internet [TAN 01a]. Sur ce point, les cartes de Mercator et de *nec* exhibent des distributions similaires, bien que celles de *nec* soient plus précises comme le montrent les coefficients de corrélation des régressions linéaires.

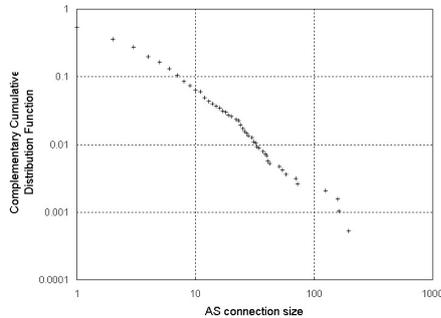


Figure 9. *Distr. cum. comp. de la taille des connexions vues par Mercator.*

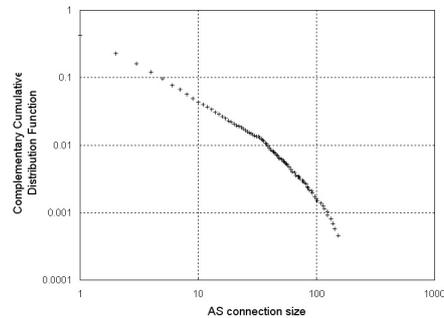


Figure 10. *Distr. cum. comp. de la taille des connexions vues par nec.*

7.4. Principaux noeuds des recouvrements

L'utilisation des recouvrements nous donne beaucoup d'information sur leurs principaux (i.e. plus gros) noeuds. Le tableau 4 contient la liste des 10 plus gros AS présents dans la table BGP. Cette liste est typique de tous les travaux sur les AS effectués depuis la mise en marche de *route-views* en 1997. Elle est utile pour l'étude des tableaux suivants.

Tableau 4. *Top 10 des AS*

Rang	no d'AS	Nom	Degré
1	701	Alternet	2502
2	1239	SprintLink	1792
3	7018	AT&T	1604
4	209	Qwest	850
5	3561	Cable&Wireless	729
6	3549	GlobalCrossing	606
7	1	Genuity	583
8	702	UUNET	579
9	3356	Level3	577
10	2914	Verio	537

Le tableau 5 contient la liste des 10 plus gros routeurs présents dans la carte créée par Mercator depuis notre laboratoire. Cette table révèle quatre points intéressants :

- Aucun routeur du top 10 des noeuds IP n'appartient à un des AS du top 10 des AS ! Ceci est très probablement dû au biais lié au point de départ de nos mesures avec Mercator (i.e. Illkirch, France).

- Un autre effet de ce biais est que le plus gros routeur (souvenons nous que d'après la section 5, ce dernier n'est pas nécessairement un routeur réel) est en Allemagne et le second en France. Encore une fois, cela s'explique par le fait que notre

14 Nom de la revue ou conférence (à définir par `\submitted` ou `\toappear`)

point de départ des mesures, Illkirch, se trouve pratiquement sur la frontière allemande.

– Beaucoup de routeurs appartiennent au réseau polonais, qui n’est pas réputé pour être l’un des plus gros du monde. Ce biais est dû à la méthode utilisée par Mercator pour sonder les cibles. Il procède par exploration des adresses consécutives dans le plan d’adressage IP. Il est clair que Mercator n’a pas eu le temps d’explorer des réseaux plus gros.

– Seulement 6 routeurs sur 10 sont BGP selon notre méthode de détermination des routeurs BGP présentée dans la section 6. L’exploration de Mercator se focalise plus sur les topologies intra-domaines de part son algorithme de ciblage aléatoire.

Tableau 5. *Top 10 des routeurs IP trouvés par Mercator*

Rang	Adresse IP	Nom du réseau	Pays	Degré	BGP	no d’AS
1	141.48.25.33	MLU-LAN	DE	409	non	680
2	193.251.252.30	OPENTRANSIT-BB	FR	335	oui	5511
3	202.8.94.2	SINGAREN	SG	198	oui	7610
4	213.248.67.90	TELIA-TIC-NET-1	EU	165	oui	1299
5	195.205.0.130	TPSA-SZK	PL	119	oui	5617
6	195.205.0.2	TPSA-SZK	PL	115	oui	5617
7	195.205.0.106	TPSA-SZK	PL	111	non	5617
8	193.2.31.100	ARNES-NET3-4	SI	109	non	2107
9	195.205.0.54	TPSA-SZK	PL	106	oui	5617
10	146.97.40.82	JANET-IP2	GB	103	non	786

Le tableau 6 contient la liste des 10 plus gros routeurs présents dans la carte créée par *nec* depuis notre laboratoire. Cette table révèle aussi quatre points intéressants :

– 3 routeurs du top 10 des noeuds IP appartiennent à des AS du top 10 des AS ce qui est plus conforme à la réalité du déploiement d’Internet.

– Le plus gros routeur est en Angleterre et le second aux USA. Cependant les degrés respectifs des 8 premiers sont très proches (de 92 pour le 8ème à 102 pour le 1er). Ce classement pourrait donc aisément varier si l’on poursuivait les mesures. Lors de campagnes de mesures de plusieurs mois, les plus gros noeuds IP ont en général plus d’un millier de liens et non une centaine. Cela est dû à la courte période de nos mesures (i.e. 10 jours).

– La majorité des routeurs sont localisés aux USA et les autres sont en Europe occidentale, ce qui encore une fois est plus conforme à la réalité du déploiement d’Internet.

– 10 routeurs sur 10 sont BGP selon notre méthode de détermination des routeurs BGP présentée dans la section 6. Le nombre très important des liens découverts par *nec* augmente la connectivité inter-domaine, d’où ce résultat.

8. Cartographie de l’Internet IPv6

Dans cette section, nous expliquons comment nous avons collecté une carte globale de l’Internet IPv6 actuel à l’aide de notre logiciel de cartographie et nous décri-

Tableau 6. *Top 10 des routeurs IP trouvés par nec*

Rang	Adresse IP	Nom du réseau	Pays	Degré	BGP	no d'AS
1	212.87.77.73	NEWNET	GB	102	oui	9191
2	154.14.65.2	PSINET-B2-14	USA	100	oui	3291
3	63.218.12.62	CAIS-CIDR7	USA	99	oui	3491
4	164.128.33.14	IPP-BACKBONE	CH	95	oui	3303
5	63.218.12.58	CAIS-CIDR7	USA	94	oui	3491
6	144.223.27.134	SPRINT-INNET	USA	93	oui	1239
7	193.23.248.43	HESSENCL	DE	92	oui	8365
8	166.63.210.62	CW-NETCS2	USA	92	oui	3561
9	166.63.210.61	CW-NETCS2	USA	91	oui	3561
10	195.8.0.94	CPRMNET	PT	84	oui	8657

vons quelques propriétés topologiques de cette carte en comparaison avec celles de l'Internet IPv4. A ce jour nous sommes à notre connaissance les premiers à présenter un tel résultat. Celui-ci permet de se faire une idée du déploiement d'IPv6 dans l'Internet actuel.

La sélection des adresses IPv6 (i.e. des cibles pour les requêtes *traceroute*) est basée sur leur préfixe. La campagne de collecte a été effectuée en deux étapes. Dans la première étape, toutes les adresses IPv6 appartiennent à un préfixe IPv6 distinct récupéré dans les bases de données [?]. Ces adresses de destinations ont été générées en concaténant le préfixe IPv6 et un identifiant Ethernet EUI-64 fictif. Pour améliorer la précision de la première carte obtenue à l'issue de la première étape et parce que les adresses IPv6 générées ont une forte probabilité de ne pas exister sur le réseau, nous avons effectué une deuxième étape de collecte en utilisant cette fois les adresses IPv6 réelles obtenues lors de la première étape de collecte *traceroute*.

8.1. Collecte de la carte Internet IPv6

Le fonctionnement de *nec* pour cartographier l'Internet IPv6 et l'Internet IPv4 est très similaire. Il existe pour cela des serveurs web *traceroute* pouvant effectuer des traces IPv6. De même nous voulons cartographier l'Internet IPv6 de la manière la plus extensive possible. A notre connaissance nous sommes les premiers à produire une carte de l'Internet IPv6 aussi récente en utilisant la technique des serveurs *traceroute* distribués. Pour collecter cette carte, nous avons lancé *nec* pendant 15 jours. Comme l'indique la table 1, nous avons effectué deux collectes pendant ces 15 jours car la liste des premières cibles était trop petite et pas assez réaliste (i.e. dû à la création d'adresses fictives). La première collecte a utilisé 628 cibles et 40 traceurs, et nous a permis de découvrir 2652 adresses IPv6 supplémentaires. La deuxième collecte a utilisé les nouvelles adresses découvertes en plus des adresses fictives. En totalité, nous avons découvert 5909 noms et/ou adresses IPv6.

De la même manière que pour la campagne de cartographie Internet IPv4, la durée de 15 jours est dû au fait que nous limitons le nombre de requêtes simultanées sur un serveur *traceroute* pour éviter de surcharger leur lien. Cette durée ne peut pas être

comparée à une durée de campagne de cartographie IPv4 mais est déjà très intéressante compte tenu du fait que nous produisons une carte la plus exhaustive possible.

Il y a très peu de filtrage ICMP sur le réseau IPv6 Internet. Ceci est dû à l'âge relativement récent de ce réseau et parce qu'il est encore considéré comme un réseau expérimental par beaucoup d'opérateurs. Cette propriété nous permet d'effectuer une résolution d'alias (i.e. désambiguation des interfaces IP) plus facilement, plus particulièrement en utilisant les options de "source-routing" ce qui s'avère impossible à faire dans le réseau IPv4.

La résolution des alias est tout aussi important en IPv6 qu'en IPv4 et peut se révéler plus difficile compte tenu de la taille de l'espace d'adressage de la nouvelle version du protocole IP. Pour résoudre les alias nous avons utilisés plusieurs méthodes qui se répartissent en méthodes négatives et positives. Les méthodes négatives sont utilisées pour trouver si deux adresses n'appartiennent pas au même noeud IPv6 afin de diminuer le nombre d'alias à résoudre par les méthodes positives qui sont généralement plus coûteuses. Nous avons utilisé trois méthodes négatives et une méthode positive de résolution des alias. Un nombre important d'interfaces découvertes contenaient seulement le nom DNS et pas l'adresse IPv6 car la commande *traceroute* IPv6 ne fournit pas systématiquement l'adresse IP lorsqu'elle arrive à résoudre le nom de domaine. A cause de cela nous avons défini une méthode négative basée sur le DNS. Celle-ci est la même que celle décrite en section 4.

La deuxième méthode négative est basée sur le fait que chaque adresse découverte par le même serveur *traceroute* est sur un arbre et qu'a fortiori toutes celles d'un même arbre ne peuvent pas être des alias potentiels. La dernière méthode négative fait l'hypothèse que chaque noeud IPv6 dispose d'une seule pile IPv6 et d'une seule table de routage. Nous effectuons un "ping" des deux adresses IPv6 et si le TTL n'est pas le même pour les deux paquets ICMP retournés, il ne sont plus considérés comme des alias potentiels.

En utilisant les méthodes négatives, nous sommes passés de n^2 couples d'alias potentiels avec $n = 5909$ à $n = 250$. Nous avons testé ces couples potentiels avec une méthode de résolution des alias positive. Cette méthode est basée sur le fait que si un message ICMP est envoyé vers une destination, il est envoyé avec l'adresse source de l'interface utilisé pour atteindre la destination. Cette méthode a été initialement proposée par Pansiot *et al.* dans [PAN 98] et a été décrite dans la section 4.

8.2. Comparaison de l'Internet IPv4 avec l'Internet IPv6

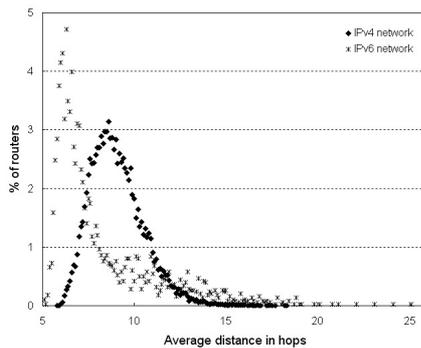
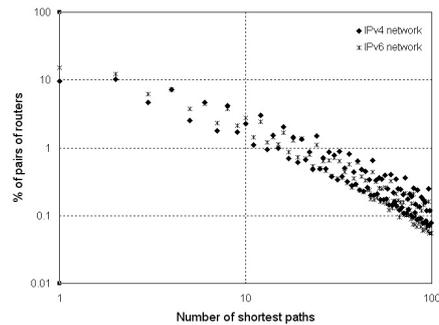
Dans cette section, nous comparons les résultats obtenus dans les sections précédentes dans lesquelles nous avons utilisé *nec* pour obtenir une carte de l'Internet IPv4 avec la carte IPv6 obtenue lors de ce travail. Nous montrons que même si le réseau IPv6 actuel est très petit comparé au réseau IPv4, il contient des propriétés très similaires. La table 7 donne une idée de la taille des deux réseaux. Le réseau Internet IPv6 est environ dix fois plus petit que celui IPv4 et contient un nombre important d'alias.

Tableau 7. *Caractéristiques des cartes Internet IPv4 et IPv6*

Source	Nombre d'interfaces	Nombre de routeurs	Degré moyen des noeuds	nombre de liens	Date
<i>nec</i> IPv4	50123	47055	5.1	119909	19-29/4/2003
<i>nec</i> IPv6	5909	3834	6.5	12401	27/5-10/6/2003

Ceci peut être expliqué par le fait que le réseau IPv6 actuel contient proportionnellement plus de tunnels que le réseau IPv4.

La figure 11 nous montre que la distance moyenne sur le réseau IPv6 est plus petite que la distance moyenne en IPv4. Les deux réseaux sont fortement maillés. La figure 12 nous montre que la distribution du nombre de plus courts chemins suit une loi puissance tout comme en IPv4. La carte IPv6 contient globalement plus de liens, ce qui augmente le nombre de plus courts chemins. Cette propriété peut être exploitée pour diminuer le nombre d'états dans les routeurs comme montré dans [HOE 03].

**Figure 11.** *Distribution de la distance moyenne***Figure 12.** *Distribution du nombre de plus courts chemins*

9. Cartographie IPv4 étendue

Notre première campagne était surtout destinée à évaluer la rapidité et la précision avec lesquelles nous pouvions cartographier l'Internet par comparaison à Mercator. Cependant, en considérant le nombre de routeurs (i.e. 47k) la carte collectée était plutôt petite. Pour notre deuxième campagne nous avons voulu obtenir une carte la plus complète possible. Nous avons voulu garder les propriétés de notre outil (i.e. découverte des liens transverses) tout en collectant une carte beaucoup plus grande. Nous avons utilisé *nec* comme dans la première campagne mais nous avons changé des paramètres et des jeux de données. Pour les adresses cibles nous avons créé des

Tableau 8. *Comparaison des campagnes*

Paramètre	1ere campagne	2eme campagne	Ratio
Date de fin	4/29/2003	9/15/2003	-
Durée (jours)	10	40	4.00
Traceurs	76	228	3.00
Cibles	6013	13828	2.30
Routeurs	47055	75885	1.61
Liens	119909	357317	2.98
Traces	456988	3152784	6.90
Trafic (traces/jour)	45698.8	78819.6	1.72
Efficacité (routeurs/trace)	0.103	0.024	0.23
Efficacité (liens/trace)	0.262	0.113	0.43

adresses aléatoires à partir de chaque préfixe d'une table de routage BGP de *route-views*, les préfixes plus long que /16 étant réduits à /16. Cela nous a donné presque 14000 adresses cibles. Nous avons ajouté de nombreux nouveaux serveurs web *traceroute* dans *nec*. La campagne a nécessité 40 jours pour être complétée ce qui est 4 fois plus long que la première campagne. Une comparaison des paramètres est donnée dans la table 8.

Si la quantité de cartographie de *nec* était linéaire et étant donné les chiffres de la table 8, nous devrions avoir collecté une carte 7 fois plus grosse que la première. Cela représente une carte de 325k routeurs avec 827k liens ! Cependant, les valeurs de la table 8 montre que cela n'a pas été le cas. L'efficacité de la cartographie a été plus faible dans la deuxième campagne et cela illustre les limites d'une approche purement quantitative. Une explication partielle de cette limitation est que de larges parts de l'Internet public ne sont pas "cartographiable" à cause du filtrage ICMP (activé pour des raisons de sécurité en général) et nous serons incapables de tracer ces zones peu importe la quantité de traceurs et de cibles que nous utilisons. Une autre explication est que notre ensemble de cibles ne soit pas optimum (i.e. nous n'observons pas tout l'espace d'adressage). Nous devons donc définir nos traces planifiées avec prudence afin d'éviter des traces redondantes qui réduisent l'efficacité de *nec* et afin d'augmenter la quantité de traces intéressantes (i.e. qui apportent beaucoup de nouveaux routeurs/liens). Nous étudions actuellement le meilleur moyen d'optimiser la quantité de traces par rapport à leur apport d'information. Entre les deux campagnes, la quantité de routeurs trouvés par trace a chuté de 77% alors que la quantité de liens par trace a diminué de 57%. Des *traceroute* incrémentaux (i.e. ne commençant pas à la source mais au milieu du chemin) permettrait de réduire grandement les traces inutiles mais cela nécessite une version nouvelle de *traceroute*. Nous devrions alors déployer nos propres traceurs et nous ne pourrions pas utiliser les serveurs *traceroute* actuels ainsi que les lentilles d'observation (i.e. "looking glass" en anglais). Nous pourrions cependant associer les deux méthodes.

9.1. Inflation des chemins

Nous avons étudié l'inflation des chemins Internet à l'aide des traces et de la carte produites par notre deuxième campagne. Tout *traceroute* définit un chemin réellement emprunté par un paquet depuis le serveur *traceroute* (source) jusqu'à l'adresse cible (destination). Après avoir créé la carte, nous pouvons calculer les plus courts chemins entre toutes les paires de routeurs en utilisant l'algorithme de Dijkstra pour chaque routeur de la carte. Nous obtenons parmi tous les plus courts chemins, tous ceux entre les sources et les destinations de nos chemins *traceroute*. Nous mesurons la longueur des chemins *traceroute* et des plus courts chemins en utilisant pour métrique le nombre de sauts. L'inflation d'un chemin est égal à la longueur d'un chemin *traceroute* d'une paire source-destination divisée par la longueur du (d'un) plus court chemin de cette même paire. Le résultat de la division est un nombre réel supérieur ou égal à 1. Nous l'appelons le ratio de la longueur du chemin ou aussi l'inflation du chemin. Avec les traces de la deuxième campagne, nous avons pu calculer le ratio de 2196960 chemins (i.e. paires source-destination).

La distribution cumulative des ratios des 2M de paires est tracée sur la figure 13. Les ratios varient de 1 à 10. Nous avons fixé 10 comme limite supérieure. Nous avons mesuré quelques valeurs supérieures à 10 mais nous les considérons anormales. Ces valeurs anormales peuvent être expliquées par le fait que les chemins *traceroute* proviennent des données brutes alors que les plus courts chemins proviennent de la carte des routeurs après désambiguation des interfaces et élimination des bouts de graphes non connexes. Donc certaines interfaces IP et liens IP ont disparu de la carte et peuvent être la cause de chemins ayant des propriétés étranges telles qu'une inflation très élevée. L'un des faits les plus intéressants de ces données est que seulement 2.72% des chemins ont un ratio de 1 (i.e. cas où le chemin *traceroute* a la même longueur que le plus court chemin correspondant). De plus, 43.4% des chemins ont un ratio supérieur à 2 ce qui est clairement non négligeable ! Cela correspond à une inflation du chemin d'au moins 100% par rapport au plus court chemin correspondant.

Ces résultats confirment que l'inflation des chemins n'est pas un phénomène marginal mais qu'il a un impact très important sur la majorité des chemins. Les causes peuvent être nombreuses :

- Cela est dû principalement aux politiques de routage dans l'Internet surtout en inter-domaine. Les liens et les routes peuvent avoir des poids/coûts de routage fixés administrativement (surtout en inter-domaine) et donc les routes peuvent ne pas être optimales. De plus dans le routage inter-domaine, il est techniquement impossible de déterminer le plus court chemin au niveau routeur d'un routeur vers un autre routeur situé dans un AS différent car seul un sous-ensemble de la topologie intra-domaine est annoncée par les protocoles de routage intra-domaines au protocole inter-domaine BGP.

- De nombreux chemins sont coupés dans les traces (i.e. ils s'arrêtent avant d'atteindre la destination désirée) à cause du filtrage ICMP. Nous utilisons la vraie destination (i.e. le dernier routeur qui répond) et non le routeur cible original pour déterminer

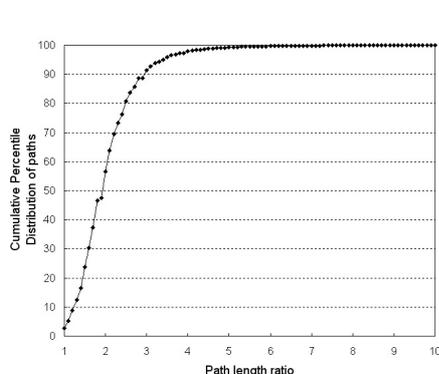


Figure 13. Distribution cumulative de l'inflation des chemins.

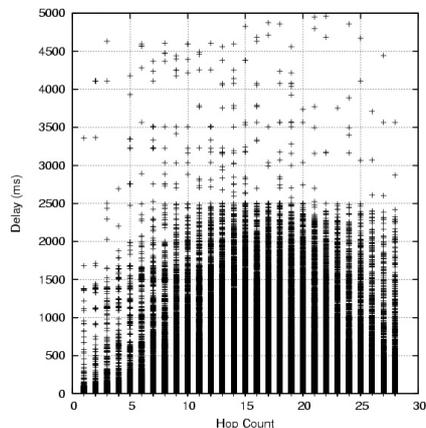


Figure 14. Délais des chemins vs nombre de sauts des chemins.

le ratio mais si nous avons des traces complètes, peut être que les ratios seraient plus faibles surtout si il n'y a pas d'inflation dans la partie manquante du chemin.

– Cela peut aussi être causé par notre sélection de destinations comme cibles *traceroute*. Nous ne choisissons peut-être pas des chemins qui reflètent des chemins typiques (i.e. entre sources et destinations plus fréquentes que d'autres).

Nous avons aussi étudié dans les traces de notre deuxième campagne, les valeurs des temps aller-retour (i.e. Round-Trip Time en anglais) et des délais. Nous avons cherché une corrélation entre la longueur d'un chemin exprimée en nombre de sauts et sa latence (ou délai, qui est égal au RTT donné par le dernier routeur ayant répondu et divisé par 2). Chaque point de la figure 14 représente un chemin avec sa longueur en abscisse et sa latence en ordonnée. 1,34M de chemins sont représentés car toutes les traces n'ont pas forcément des informations sur le RTT qui sont exploitables. On ne peut observer sur la figure ni linéarité, ni corrélation particulière. Le seul phénomène que l'on peut voir est une borne maximum du délai du chemin en fonction de sa longueur (en écartant les points extrêmes). Sous cette borne, on trouve toutes les valeurs de délais possibles. Le tracé suivi par l'évolution de la valeur de cette borne maximum ressemble un peu à une distribution normale. On s'attendrait à ce que cette borne augmente régulièrement avec la longueur des chemins, mais nous devons prendre en compte que les valeurs RTT mesurées par les *traceroute* sont dynamiques et dépendent non seulement de la longueur du chemin (délai théorique minimum) mais aussi de la charge de certains routeurs du réseau (engorgement, files d'attente avec diverses priorités, etc.). Donc il faudrait différencier nos mesures entre les délais minimaux (les paquets des traces n'ont subi aucun ralentissement) et les délais dynamiques réels (les paquets des traces ont été ralentis à un ou plusieurs endroits) : cela est tech-

niquement impossible à réaliser avec les seules informations fournies par les serveurs *traceroute*. Le fait que la part relative des délais minimaux (traces non freinées) soit plus importante dans les longs chemins (20 sauts ou plus) peut expliquer le fait que la valeur de la borne maximale soit plus faible que pour des chemins plus courts.

L'inflation des chemins a été étudiée par Tangmunarunkit *et al.* dans [TAN 01b] en 2000. Ils ont trouvé que plus de 95% de leurs chemins ont un ratio inférieur ou égal à 2 alors que nous n'en avons trouvés que 56.6%. Cependant leur étude ne portait que sur 61k traces alors que nous en avons 2M et ils ont utilisé Mercator pour créer la carte (i.e. donc cette dernière manquait un grand nombre de liens transverses, voir section 7). C'est pourquoi nous pensons que nos résultats sont plus précis. Notons qu'une augmentation du nombre de sauts due à l'inflation de chemin entre deux routeurs peut ne pas se traduire automatiquement par une forte augmentation du délai sur ce chemin. Etudier l'inflation des chemins en utilisant le délai comme métrique fait partie de nos objectifs ultérieurs. Très récemment, Spring *et al.* ont effectué une étude proche de la nôtre dans [SPR 03] et ils ont trouvé que l'augmentation moyenne du délai des chemins pour les chemins inter-domaines (le plus mauvais cas) est inférieur à 14 ms. Cependant ils n'ont pas étudié la corrélation entre l'inflation du délai et celle du nombre de sauts.

9.2. Etude des délais

Comme cela est précisé dans la section précédente, nous avons aussi étudié dans les 3Go de traces de notre deuxième campagne, les valeurs des RTT donnés par les *traceroute*. Ci-dessous l'exemple d'une trace typique :

```
traceroute to 24.100.0.103 (24.100.0.103) from ns2.acadia.net (142.167.1.1), 40 hops max, 40 byte packets
 1 ip142167001252.acadia.net (142.167.1.252) 0.677 ms 0.622 ms 0.481 ms
 2 63.111.121.197 (63.111.121.197) 8.676 ms 8.501 ms 9.339 ms
 3 196.at-1-0-0.CL2.BOS1.ALTER.NET (152.63.25.126) 8.682 ms 8.678 ms 9.095 ms
 4 0.so-6-2-0.XL2.NYC9.ALTER.NET (152.63.29.114) 15.170 ms 15.130 ms 15.149 ms
 5 POS7-0.BR1.NYC9.ALTER.NET (152.63.18.221) 15.228 ms 15.087 ms 15.244 ms
 6 acr2-so-5-3-0.NewYork.cw.net (206.24.193.245) 16.759 ms 16.170 ms 15.929 ms
 7 agr4-loopback.NewYork.cw.net (206.24.194.104) 16.613 ms 16.955 ms 16.538 ms
 8 dcr2-so-7-3-0.NewYork.cw.net (206.24.207.205) 15.753 ms 16.104 ms 15.798 ms
 9 dcr1-loopback.Chicago.cw.net (208.172.2.99) 38.989 ms 39.370 ms 39.279 ms
10 telus-rogers.Chicago.cw.net (208.175.10.86) 49.834 ms 49.704 ms 47.761 ms
11 gw01.esna.phub.net.cable.rogers.com (66.185.80.193) 50.105 ms 50.021 ms 50.653 ms
12 gw01.mtmc.phub.net.cable.rogers.com (66.185.81.166) 48.114 ms 47.789 ms 50.860 ms
13 gw01.ajax.phub.net.cable.rogers.com (66.185.81.170) 50.717 ms 50.375 ms 50.563 ms
14 66.185.81.198 (66.185.81.198) 50.734 ms 50.530 ms 50.718 ms
15 tlgw1.ajax.phub.net.cable.rogers.com (24.100.0.1) 50.686 ms 51.066 ms 51.040 ms
16 * * *
17 * * *
```

Pour une ligne donnée, les valeurs temporelles exprimées en millisecondes, représentent le temps aller retour d'un paquet depuis la source (le serveur *traceroute*) jusqu'au routeur dont l'adresse est donnée en début de ligne. Trois paquets sont envoyés afin d'éviter des valeurs biaisées par un retard de traitement momentané sur le trajet d'un paquet et donc trois valeurs sont mesurées. Ces valeurs permettent de déduire les délais sur les liens de la façon suivante :

1) On calcule la moyenne m_n des trois valeurs de RTT pour le routeur n (on pourrait envisager un traitement plus complexe que la simple moyenne).

- 2) On calcule la moyenne m_{n-1} des trois valeurs de RTT pour le routeur $n - 1$.
- 3) On calcule $(m_n - m_{n-1})/2$ pour obtenir le délai du lien de $n - 1$ vers n .

Il arrive que le résultat soit négatif (avec notre méthode, le délai entre 11 et 12 ci-dessus sera négatif). Dans ce cas, si il est négatif et peu important (proche de zéro), alors on attribue au lien un délai de 1 ms. A noter que les traces peuvent passer de très nombreuses fois sur un même lien. On refait le calcul à chaque passage sur un lien donné et on stocke toutes les valeurs dans un tableau qui lui est propre. Lors de la construction de la carte, on attribue à chaque lien un délai égal à la moyenne de tous les délais stockés dans son tableau. A noter que pour les délais, on considère les liens comme asymétriques. Donc on calcule les délais en tenant compte du sens de parcours du lien et par conséquent chaque lien a donc deux délais : le délai du routeur $n - 1$ vers n et celui du routeur n vers $n - 1$. Toutes les traces n'ont pas toujours des informations de RTT exploitables (pour des raisons techniques), c'est pourquoi la carte que nous avons créé qui contient des valeurs de délais pour les liens est légèrement plus petite que celle qui ne contient que des informations de connectivité même si les deux ont été réalisées à partir des mêmes traces de notre deuxième campagne. La carte Internet avec les délais des liens est disponible sur le site de *nec* [MAG a]. La carte avec les délais contient 281957 liens (contre 357317 pour la carte simple). Cela nous donne 563914 valeurs de délais à déterminer (2 par lien). Nous attribuons la valeur 0 à tout délai de lien que nous n'avons pas pu déterminer. En effet les traces ont un sens de parcours et comme il y a beaucoup plus de destinations que de sources, la majorité des trajets ne sont jamais fait en sens inverse. Notre technique nous a permis de déterminer les valeurs de 194431 délais, soit un taux détermination de 34,4%. Les autres délais ont pour valeur 0 (i.e. indéterminé).

La figure 15 illustre la distribution de ces 194k délais. Celle-ci est fortement biaisée et sa partie de gauche pourrait presque être approximée par une loi puissance. La majorité des liens (considérés dans le sens de leur mesure) ont donc des délais très faibles, ce qui est logique. La distribution cumulative de ces mêmes délais illustrée sur la figure 16 fait apparaître un tracé plus compact, mais pas une linéarité. On ne peut donc affirmer se trouver devant une distribution à queue lourde bien que le tracé s'en rapproche, cependant n'oublions pas que ces délais sont très dynamiques (ils dépendent de l'état du réseau lors de leur mesure). On peut noter deux points intéressants sur la figure 16 : seulement 28,5% des liens ont un délai supérieur à 10 ms et moins de 1,8% des liens ont un délai dépassant les 100ms. On peut raisonnablement considérer que les liens ayant des délais supérieurs à 100ms sont soit de très longue connexions transcontinentales, soient des liaisons satellites, soient que leurs mesures ont été faussées par des engorgements du réseau.

La faible corrélation des délais des chemins et de leurs longueurs en nombre de sauts est un résultat nouveau à notre connaissance et très intéressant. Cela signifie que les simulations de protocoles ne doivent plus se contenter de délais identiques sur tous les liens comme cela se voit encore très fréquemment dans les topologies des simulations utilisant *ns-2* par exemple. En effet cela revient à utiliser la métrique du nombre de sauts. Nos résultats de la figure 14 montrent que le nombre de sauts

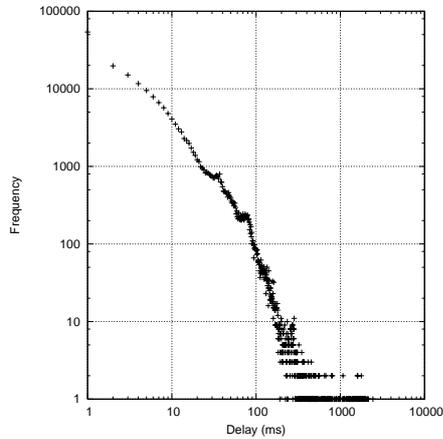


Figure 15. *Distribution des délais des liens.*

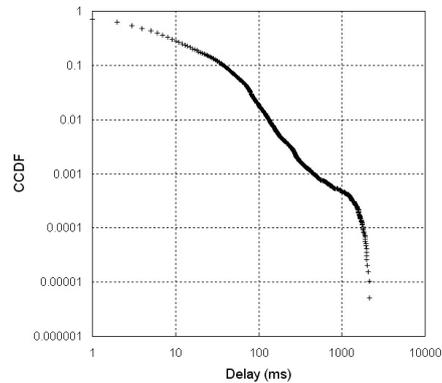


Figure 16. *Distribution complémentaire cumulative des délais des liens.*

d'un chemin n'est pas directement corrélé au délai de celui-ci. Il est donc préférable d'utiliser des délais différents sur les liens et dont les valeurs se rapprochent de nos résultats. Les topologies utilisées dans les simulations de protocoles Internet doivent donc s'inspirer de la distribution des délais des liens de la figure 15 afin d'accroître leur réalisme.

Nous avons cherché à découvrir une corrélation entre la taille d'un routeur exprimée par le nombre de liens IP adjacents sortants à celui-ci et le délai de ses mêmes liens IP adjacents sortants. Chaque point de la figure 17 représente un lien (dans son sens de mesure) avec son délai en abscisse et le degré de son routeur de départ en ordonnées. L'idée étant que plus un routeur est interconnecté (degré élevé), plus il est en coeur de réseau et donc plus il possède de connexions rapides. Le tracé nous montre que cette idée n'est pas fondée. Il n'apparaît pas de corrélation évidente entre la taille d'un routeur et le délais de ses liens adjacents sortants. De même nous avons cherché à découvrir une corrélation entre la position d'un routeur exprimée par sa distance moyenne à tous les autres routeurs mesurée en nombre de sauts et le délai de ses mêmes liens IP adjacents sortants. Chaque point de la figure 18 représente un lien (dans son sens de mesure) avec son délai en abscisse et la distance moyenne de son routeur de départ en ordonnées. L'idée étant que plus un routeur est proche du centre du réseau (distance moyenne faible), plus il possède de connexions rapides. Encore une fois, il n'apparaît pas de corrélation évidente entre la position d'un routeur et le délais de ses liens adjacents sortants.

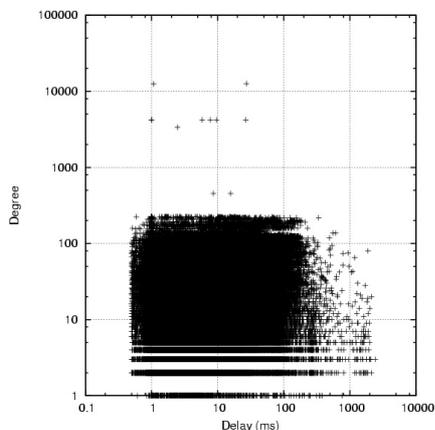


Figure 17. Degré des routeurs vs délais des liens.

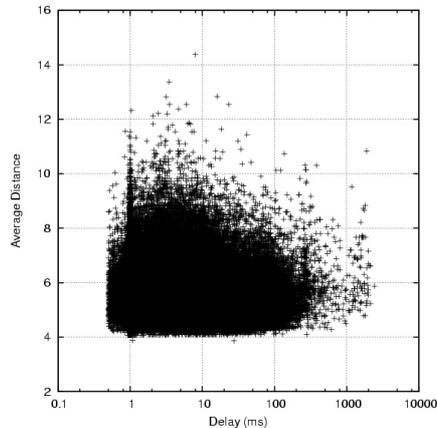


Figure 18. Distance moyenne des routeurs vs délais des liens.

10. Conclusion

La cartographie d'Internet est une opération complexe. Dans cet article nous avons décrit notre logiciel *nec* de cartographie rapide et nous avons montré que celui-ci est nettement supérieur à Mercator, seul logiciel libre de cartographie actuellement disponible. En 10 jours, *nec* a découvert presque autant de routeurs et surtout presque deux fois plus de liens que Mercator. La carte obtenue possède une quantité relative de liens supérieure à une carte d'AS ce qui est un résultat nouveau pour une carte de niveau routeur à l'échelle de l'Internet public. De plus les recouvrements nous montrent que *nec* produit une carte beaucoup moins biaisée que Mercator. Nous avons montré que *nec* peut aussi cartographier la topologie IPv6 et que ce dernier bien que très petit, partage certaines caractéristiques du réseau IPv4. Cela est logique étant donné le fait que ce réseau se déploie le plus souvent sur du matériel appartenant au réseau IPv4. Nous avons aussi évalué la capacité de cartographie maximale de notre logiciel et avons montré que celle-ci atteint des limites liées à son architecture et à celle de l'Internet lui-même (filtrage intensif). Nous avons enfin étudié les délais des chemins et des liens de l'Internet IPv4. Nous avons observé que la longueur en nombre de sauts des chemins est faiblement corrélée au délai de ces mêmes chemins. Nous avons découvert que la distribution des liens est fortement biaisée et se rapproche d'une distribution à queue lourde ce qui est là encore un résultat inédit à notre connaissance. Plus de deux tiers des liens de l'Internet ont un délai mesuré inférieur à 10ms. Nous travaillons actuellement à l'optimisation du nombre et de l'orientation des traces nécessaires à mener à bien les cartographies ainsi qu'à l'amélioration de la détermination des délais mesurés sur les liens. En effet, les délais ont une nature hautement dynamique liée à l'activité du réseau. Nous espérons que les résultats produits par notre logiciel (cartes

IPv4 et IPv6, recouvrements routeurs-AS, délais des liens, etc.) seront profitables à la communauté réseau pour les simulations faisant appel à une modélisation réaliste de l'Internet. Les personnes désireuses d'utiliser les cartes produites par *nec* (et contenant les recouvrements associés) pour faire des simulations de protocoles peuvent les trouver à l'adresse [MAG a].

11. Bibliographie

- [BAR 01] BARFORD P., BESTAVROS A., BYERS J., CROVELLA M., « On the Marginal Utility of Network Topology Measurements », *Proceedings of ACM SIGCOMM Internet Measurement Workshop'01*, November 2001, p. 5-17.
- [BU 02a] BU T., GAO L., TOWSLEY D., « On characterizing BGP routing table growth », *Proceedings of IEEE GLOBECOM'02*, 2002, p. 2197-2201.
- [BU 02b] BU T., TOWSLEY D., « On distinguishing between Internet power law topology generators », *Proceedings of IEEE INFOCOM'02*, New York City, NY, USA, June 2002.
- [BUR 99] BURCH H., CHESWICK B., « Mapping the Internet », *IEEE Computer*, vol. 32, n° 4, 1999, p. 97-98.
- [CHA 01] CHANG H., JAMIN S., WILLINGER W., « Inferring AS-Level Internet Topology from Router-Level Path Traces », *Proceedings of SPIE ITCOM'01*, Denver, CO, USA, August 2001.
- [CHA 02] CHANG H., GOVINDAN R., JAMIN S., SHENKER S., WILLINGER W., « Towards capturing representative AS-level Internet topologies », rapport, 2002, University of Michigan.
- [CHE 02] CHEN Q., CHANG H., GOVINDAN R., JAMIN S., SHENKER S., WILLINGER W., « The origin of power laws in Internet topologies revisited », *Proceedings of IEEE INFOCOM'02*, New York City, NY, USA, June 2002.
- [GOV 00] GOVINDAN R., TANGMUNARUNKIT H., « Heuristics for Internet Map Discovery », *Proceedings of IEEE INFOCOM'00*, Tel Aviv, Israël, March 2000.
- [HOE 03] HOERDT M., PANSIOT J.-J., MAGONI D., GRAD D., « Répartition des états de routage multipoint par les plus courts chemins multiples », *Actes des 5èmes rencontres francophones sur les aspects Algorithmiques des Télécommunications*, Banyuls-sur-mer, France, Mai 2003, p. 139-145.
- [MAG a] MAGONI D., « network cartographer (nec) », Université Louis Pasteur, <http://www-r2.u-strasbg.fr/~magoni/nec/>.
- [MAG b] MAGONI D., « network manipulator (nem) », Université Louis Pasteur, <http://www-r2.u-strasbg.fr/nem/>.
- [MAG 01] MAGONI D., PANSIOT J.-J., « Analysis of the autonomous system network topology », *ACM Computer Communication Review*, vol. 31, n° 3, 2001, p. 26-37.
- [MAG 02a] MAGONI D., PANSIOT J.-J., « Evaluation of Internet topology generators by power law and distance indicators », *Proceedings of the 10th IEEE International Conference On Networks*, Singapore, August 2002, p. 401-406.
- [MAG 02b] MAGONI D., PANSIOT J.-J., « Internet topology modeler based on map sampling », *Proceedings of the 7th IEEE Symposium on Computers and Communications*, Giardini Naxos, Sicily, Italy, July 2002, p. 1021-1027.

- [PAN 98] PANSIOT J.-J., GRAD D., « On Routes and Multicast Trees in the Internet », *ACM Computer Communication Review*, vol. 28, n° 1, 1998, p. 41-50.
- [REK 95] REKHTER Y., LI T., « A Border Gateway Protocol 4 (BGP-4) », Request For Comments n° 1771, March 1995, Internet Engineering Task Force.
- [SPR 02] SPRING N., MAHAJAN R., WETHERALL D., « Measuring ISP Topologies with Rocketfuel », *Proceedings of ACM SIGCOMM'02*, Pittsburgh, PA, USA, August 2002.
- [SPR 03] SPRING N., MAHAJAN R., ANDERSON T., « Quantifying the Causes of Path Inflation », *Proceedings of the ACM SIGCOMM'03*, August 2003.
- [TAN 01a] TANGMUNARUNKIT H., DOYLE J., GOVINDAN R., JAMIN S., SHENKER S., WILLINGER W., « Does AS size determine degree in AS topology ? », *ACM Computer Communication Review*, vol. 31, 2001.
- [TAN 01b] TANGMUNARUNKIT H., GOVINDAN R., SHENKER S., ESTRIN D., « The Impact of Routing Policy on Internet Paths », *Proceedings of IEEE INFOCOM'01*, Anchorage, Alaska, USA, 2001.
- [Uni] University of Oregon, Advanced Network Technology Center, <http://www.routeviews.org/>, « BGP data from route-views ».