Origine des étoiles dans traceroute

Farah Layouni, Brice Augustin, Timur Friedman, Renata Teixeira

Origine des étoiles dans traceroute

Farah Layouni, Brice Augustin, Timur Friedman, Renata Teixeira
Université Pierre et Marie Curie, Laboratoire LIP6–CNRS

Laboratoire LIP6
104 avenue du président Kennedy 75016 Paris
{prenom.nom}@lip6.fr

RÉSUMÉ. *Traceroute permet de découvrir le chemin entre deux machines sur internet, en envoyant des sondes avec une durée de vie limitée pour forcer les routeurs à révéler leur présence. Lorsqu'il ne parvient pas à déterminer l'adresse IP d'un routeur, traceroute affiche à la place une étoile (*). Les nombreuses étoiles présentes dans les traces introduisent des imprécisions indésirables dans les applications de cartographie de l'internet, ou encore lors d'un diagnostic de panne. Cet article se penche sur les raisons de leur apparition, que nous classons en deux catégories. (1) Le routeur émet une réponse altérée qui empêche traceroute de l'associer à une sonde émise. (2) Le routeur n'émet tout simplement pas de réponse. Nous proposons des mesures qui mettent en évidence l'existence de chacune de ces causes. De plus, nous montrons comment réduire l'effet de ces étoiles grâce à un paramétrage approprié de l'outil de traçage.*

ABSTRACT. *Traceroute is a tool to measure the route between two machines in the internet. It sends TTL-limited probes to force routers to send error messages and reveal their presence. However, it frequently fails at discovering the routers' address at some portion of the path, and prints a star (*) instead. Those stars can appear for many reasons, and we classify them into two categories. (1) The router sends a corrupted response, which prevents traceroute to assign it to a probe it sent. (2) The router simply does not send any response. We propose measurements to characterize all stars in traces collected from our university. We also show that simple solutions can help preventing some of these causes, like an appropriate setup of the tool, or a simple modification of its mechanism.*

MOTS-CLÉS : *traceroute, mesures, internet, étoiles*

KEYWORDS: *traceroute, measurements, internet, stars*

1. Introduction

Traceroute est un outil pour mesurer le chemin entre deux hôtes d'un réseau IP. Il est largement utilisé dans le diagnostic de pannes, l'inférence de propriétés et la construction de cartes de l'internet. Il découvre les routeurs sur le chemin en émettant une série de paquets avec une durée de vie croissante. Cette durée de vie ("Time to Live", ou TTL) représente le nombre maximum de routeurs que le paquet peut traverser avant d'être détruit. Lorsqu'un routeur reçoit un paquet IP, il commence par décrémenter le TTL. S'il est inférieur à 1, alors le paquet a expiré et doit être détruit. Le routeur notifie l'émetteur du paquet rejeté via un message ICMP "TTL expiré" avec comme adresse source, l'adresse d'une de ses interfaces. Traceroute envoie des paquets (appelés sondes) à faible TTL pour forcer les routeurs à émettre ce type de message et ainsi révéler leur présence.

Il arrive fréquemment que traceroute ne reçoive pas de réponse de certains routeurs, et par conséquent ne parvienne pas à déterminer leur adresse. A la place il affiche une étoile (*), dont la présence pose de nombreux problèmes. Par exemple, elles empêchent un opérateur réseau de localiser précisément une panne ; elles introduisent des erreurs lors de la construction de cartes du réseau, certaines zones restant indéterminées ; elles allongent considérablement le temps de traçage, traceroute devant attendre l'expiration d'un temporisateur avant de continuer son exploration.

Cet article est la première étude systématique des causes de ces étoiles. Nous proposons des expériences pour mettre en évidence leur existence. Nous montrons que ces causes sont nombreuses et peuvent être classées en deux catégories. Premièrement, la réponse peut parvenir à l'émetteur, mais celui-ci la rejette parce qu'il est incapable de l'associer à une sonde. Cet échec d'association vient du fait que traceroute–et les outils dérivés–utilisent diverses méthodes ad hoc, non standardisées pour identifier les sondes qu'ils envoient. Nous discutons des différents types d'altération des réponses à l'origine de ce problème dans la section 2. Deuxièmement, certains équipements n'émettent tout simplement pas de réponse. Dans la section 3 nous montrons qu'il existe une grande variété de raisons pour lesquelles un routeur ou un hôte déciderait de rester muet. Enfin, nous proposons des bilans qui montrent comment utiliser ces résultats pour mieux paramétrer nos outils de traçage.

Nous proposons des expériences pour identifier chaque phénomène mis en cause. Ces expériences impliquent des protocoles de mesure différents, que nous décrirons au moment opportun. Néanmoins, nous avons utilisé une liste de 5000 adresses générées aléatoirement et répondant au ping au moment de la construction de la liste. Nos mesures consistent généralement à lancer un traceroute vers chaque destination, à partir d'une unique source située au Laboratoire d'Informatique de Paris 6, et à faire varier certains paramètres de mesure. Nos traces couvrent 1077 systèmes autonomes, ce qui représente une très faible portion du nombre total actuel. Néanmoins nous traversons la plupart des réseaux qui composent le coeur de l'internet, avec tous les tier-1 et 64 des 100 plus gros réseaux mondiaux Notre but n'est pas d'obtenir des statistiques représentatives de l'internet actuel, sur la prépondérance de chaque cause d'étoile. Nous

cherchons plutôt à montrer leur existence et proposer des expériences qui les mettent en évidence.

2. Échecs d'associations des réponses

Traceroute doit associer chaque réponse à une sonde émise. Pour cela il place un identifiant unique dans une partie de la sonde qui est recopiée dans le message d'erreur ICMP. Traceroute opère ensuite un tri sur tous les paquets ICMP reçus, pour déterminer les réponses qui lui sont destinées. Traceroute classique utilise les numéros de ports, technique rendue obsolète par le déploiement de répartiteurs de charge par flot dans l'internet [AUG 06]. Les contraintes sont donc nombreuses pour choisir le champ approprié où placer cet identifiant, et chaque outil utilise sa propre technique. La phase d'association de réponses est une phase critique lors du traçage, car ce champ peut subir des altération lors de son traitement par des noeuds du réseau. Si l'outil n'est pas capable d'identifier une réponse lui appartenant, il peut rejeter une réponse valide (ce qui cause l'apparition d'une étoile), et manque la possibilité de découvrir un nouveau noeud dans le réseau. Il peut aussi accepter une réponse invalide, d'où l'inférence de faux liens et le calcul de délais erronés.

L'expérience menée dans cette section permet de révéler ces altérations. Nous traçons nos 5000 destinations en ICMP, UDP et TCP, et capturons tous les paquets avec `tcpdump`. Nous analysons ensuite cette trace et associons chaque réponse avec une sonde. Comme certains champs peuvent être altérés, nous associons la réponse avec la sonde qui a le plus grand nombre de champs identiques : numéros de ports, identifiant IP, checksum UDP, séquence ICMP, séquence TCP. Cette association lâche nous permet de détecter tout type d'altération dans les réponses, que nous reportons ici.

2.1. *Inversion ou réécriture de l'identifiant IP*

Pour communiquer à travers Internet, deux machines doivent utiliser un format d'encodage précis, le "Network Byte Order", qui définit l'ordre dans lequel sont représentés les octets dans le réseau. Cet ordre peut être différent de celui utilisé en interne par une machine, c'est pourquoi chaque machine doit éventuellement effectuer une conversion avant d'émettre ou recevoir des données. Certaines implémentations inversent les octets de l'identifiant IP à la réception d'un paquet, puis l'incluent ainsi inversé dans un message d'erreur ICMP envoyé à la source [MAL 07]. Par conséquent, la source reçoit une réponse comportant un identifiant erroné. Tcptraceroute utilisant ce champ, il est incapable d'associer des réponses ainsi altérées. Dans nos traces nous avons détecté 251 réponses avec un identifiant inversé. Le traçage du chemin vers 31 destinations était affecté par ce problème. Plus rare encore (un unique cas dans nos traces), une seconde implémentation effectue une copie incorrecte du paquet IP en erreur, d'où un identifiant IP contenant des données aléatoires ou remises à zéro.

2.2. *Réécriture du checksum UDP*

La somme de contrôle UDP est un champ de 16 bits utilisé pour vérifier l'intégrité de l'entête et des données UDP. Certaines implémentations modifient ce champ avant de l'inclure dans un message d'erreur ICMP. Dans nos traces, 716 réponses comportent un checksum incorrect, soit 239 chemins affectés par ce problème. Ce problème étant largement répandu, l'utilisation de ce champ est à éviter (c'est le cas de Paris traceroute, qui l'utilise pour identifier les réponses aux sondes UDP).

2.3. *Réécriture de l'adresse destination*

Le passage de passerelles effectuant une translation d'adresse altère les paquets, et a un impact sur l'association des réponses. En effet, certaines passerelles modifient l'adresse destination des paquets entrants, avant de les transférer à une machine du réseau interne, mais oublient de faire la modification inverse dans les paquets d'erreur ICMP. En conséquence, l'outil de traçage reçoit une réponse altérée ; l'adresse destination du paquet original inclus dans la réponse (l'adresse d'une machine privée) est différente de l'adresse destination utilisée lors de l'émission de la sonde (adresse publique de la passerelle NAT). A noter que les adresses trouvées dans ces réponses altérées appartiennent principalement à des blocs d'adresses privées (192.168.0.0/16 et 10.0.0.0/8), ce qui confirme la présence d'une passerelle NAT masquant la structure du réseau interne.

Nos traces ont révélé 41 réponses avec une adresse originale altérée, soit 12 chemins affectés par ce problème. Paris traceroute vérifie la consistance de cette adresse avant d'accepter une réponse et rejette toute réponse avec une adresse différente de l'adresse originale. En conséquence, il rejette certaines réponses parfaitement valides.

Bilan : association des réponses. Il semble difficile d'élaborer une technique d'association résistante aux altérations et respectant toutes les contraintes. Tcptraceroute souffre du problème de l'altération de l'identifiant IP, ce qui n'est pas le cas de Paris traceroute pour TCP (utilisation du numéro séquence TCP) et ICMP (séquence et identifiant ICMP). Une solution pour utiliser sans risque l'identifiant IP est de n'utiliser que la moitié des identifiants disponibles, de manière à empêcher l'utilisation simultanée de deux identifiants qui pourraient être confondus s'ils étaient inversés (par exemple, 0x00ff et 0xff00).

2.4. *RTT très élevés*

Tous les outils de traçage attendent une réponse pendant quelques secondes puis abandonnent. Une attente plus longue semble inutile, dans la mesure où un temps d'aller-retour de plus de 5 secondes semble très improbable. La traversée de liens transocéaniques, ou encore la faible priorité accordée au trafic ICMP dans les routeurs modernes ne peut expliquer des délais de plusieurs secondes [CHO 04]. L'analyse

des traces de paquets nous a permis de calculer les temps d'aller-retour de toutes les réponses, y compris celles éventuellement arrivées en retard et donc manquées par l'outil de traçage. Sur 263566 réponses, 99.98% sont arrivées moins de 2 secondes après l'envoi de la sonde. 59 réponses sont arrivées entre 2 et 6 secondes après l'envoi. Par défaut, le délai d'abandon de Paris traceroute est de 2 secondes, ce qui lui ferait manquer toutes ces réponses, donnant au total 22 chemins affectés par ce problème. Notons que la majorité des réponses à délai élevé provient des extrémités du réseau. 70% sont des messages "Host unreachable" ou "Port Unreachable". Les premiers présentent souvent des délais d'environ 3 secondes, ce qui peut correspondre au délai d'attente, par une passerelle d'accès à Internet, d'une réponse ARP sur un LAN.

Plus inattendu, une petit nombre de réponses présentent des RTT anormalement élevés, jusqu'à 37 secondes. Plusieurs hypothèses peuvent expliquer ces délais. Premièrement, il est possible que les sondes déclenchent l'ouverture d'une connexion pour atteindre une partie du réseau. Cette hypothèse semble à exclure puisque dans un tel cas les sondes suivantes et passant par ce chemin seraient aussi retardées ; dans nos observations, les réponses pour les sauts suivants sont arrivées avec un délai normal (moins de 200 ms), ce qui indique qu'elles ont traversé sans délai le routeur incriminé et atteint les routeurs suivants. Nous pouvons aussi envisager que les sondes ou les réponses ont été retenues dans le réseau, par exemple à cause d'une boucle de routage intermittente. Les TTL des réponses indiquent qu'elles n'ont pas traversé un nombre anormalement élevé de routeur (environ 15 dans les exemples observés), ce qui infirme cette deuxième hypothèse. Enfin, il reste l'hypothèse de la surcharge temporaire de plan de contrôle du routeur incriminé. Ce plan de contrôle est responsable des calculs réalisés de manière régulière (calcul des tables de routage, gestion de l'interface, envoi des messages d'erreur). Ce travail est généralement réalisé par un processeur généraliste, contrairement au transfert des paquets qui nécessite un matériel dédié pour atteindre les débits requis par les réseaux actuels. Il est donc envisageable que ce processeur soit momentanément surchargé à cause d'une activité qui bloque tout autre calcul, ce qui retarde la prise en charge des messages d'erreur ICMP. Plusieurs observations semblent confirmer cette hypothèse : les réponses des routeurs incriminés ne sont pas systématiquement retardées. Pendant certaines périodes, les RTT sont normaux, correspondant à des périodes où le routeur est disponible. Lorsqu'elles sont retardées, les réponses arrivent dans l'ordre dans lequel les sondes ont été émises, mais avec un délai inter-réponse très atténué par rapport au délai inter-sonde : par exemple, pour un routeur observé, il n'y a qu'un délai de 13 ms entre les 3 réponses alors que les 3 sondes sont envoyées avec un délai de 50 ms entre chaque. Cela tend à montrer que les réponses ont été bloquées puis réémises en un temps très court.

Bilan : délai d'attente des réponses. Nos résultats montrent qu'une temporisation supérieure à 2 secondes dans les outils de traçage n'apporte pas d'avantage significatif. De plus le temps d'attente des réponses influe sur la durée de la mesure. Par exemple, le temps de traçage moyen d'un chemin est de 12s avec une temporisation de 2s, et passe à 22s avec une temporisation de 5s.

Temps inter-sonde (ms)	50	250	500
Sondes sans réponse	5.5%	4.4%	3.4%

Tableau 1. *Impact du temps inter-sonde sur le pourcentage des sondes sans réponse (étoiles)*

3. Échecs de génération des réponses

Les standards divergent légèrement sur le comportement à adopter lorsqu'un routeur jette un paquet avec un TTL trop petit. La RFC 792 [POS 81] (1981) indique qu'un message d'erreur *peut* être envoyé à la source. En revanche, la RFC 1812 [BAK 95] (1995) est plus claire : tout routeur IPv4 *doit* signaler le problème à la source. La réalité est plus complexe, à cause des différentes implémentations, configurations et politiques de sécurité employées par les opérateurs. Cette section énumère les raisons pour lesquelles un routeur peut décider de jeter silencieusement un paquet sans générer le message d'erreur utilisé par traceroute.

3.1. *Rate-limiting*

Les standards prévoient qu'un routeur envoie un message d'erreur pour chaque paquet jeté, mais ne tiennent pas compte des problèmes de sécurité que cela implique (déni de service, en particulier). Pour palier ce problème certains routeurs offrent la possibilité de limiter le nombre de réponses qu'ils émettent dans un intervalle de temps donné (*rate-limiting*).

La détection des routeurs ayant activé cette option est une tâche délicate et difficile à discerner des autres phénomènes comme la perte de paquets. Pour cette raison, nous avons plutôt cherché à détecter ce phénomène au niveau global, c'est à dire son impact sur le nombre total de sondes sans réponse. Pour cela nous avons réalisé l'expérience suivante. Nous traçons la route vers nos 5000 destinations en envoyant 10 sondes successives à chaque routeur. Nous répétons l'expérience pour plusieurs temps d'attente entre les sondes destinées à un routeur : 50, 250, 500 ms. Des sondes émises en rafale (temps inter-sonde faible) et destinées à un même routeur ont plus de chance de déclencher le rate-limiting, avec pour résultat une plus grande quantité d'étoiles dans la trace obtenue.

Nous avons effectué 5 passes successives pour chaque temps inter-sonde. Dans chaque trace obtenue nous éliminons tous les sauts sans réponse qui peuvent éventuellement apparaître à la fin de la route (causés par une destination muette ou un équipement de filtrage à l'entrée du réseau de la destination, voir section 3.3). Le tableau 1 présente le pourcentage de sondes sans réponse pour les trois temps inter-sondes. Il est clair qu'un temps inter-sonde court a pour conséquence un plus grand nombre d'étoiles, confirmant la présence de rate-limiting sur certains chemins me-

surés. Nous travaillons actuellement sur une méthode pour détecter précisément les routeurs responsables de ce phénomène.

3.2. *Blocage*

Un problème important de la cartographie de l'internet concerne les routeurs dits anonymes [YAO 03] car ils bloquent toute émission de message ICMP. Par conséquent ils ne révèlent jamais leur présence et apparaissent systématiquement sous la forme d'une étoile. Plusieurs raisons peuvent expliquer la présence de tels routeurs. Par exemple, une implémentation incorrecte ou ne respectant pas la RFC 1812. Ou encore des opérateurs qui ne veulent pas révéler la structure de leur réseau, et désactivent, par une règle de filtrage, l'émission de messages ICMP par leurs routeurs.

Pour mettre en évidence ces routeurs "muets", nous devons les différencier des autres phénomènes causant des étoiles. Pour cela nous nous appuyons sur le caractère permanent des non-réponses. En effet, nous considérons que les autres cas de non-réponse ne sont que temporaires, de l'ordre de la seconde. Cette hypothèse est valable pour le rate-limiting et la perte de paquets. Par conséquent, si nous sondons un routeur pendant un laps de temps suffisamment long, nous diminuons les chances de confondre une non-réponse temporaire avec un cas de non-réponse permanente.

Nous avons donc réalisé l'expérience suivante. Nous traçons le chemin vers chaque destination en émettant 10 sondes par saut, avec une pause de 500 ms entre chaque sonde. Nous sondons donc chaque routeur pendant 5 secondes. Nous éliminons les sauts sans réponses qui peuvent apparaître à la fin des chemins. A partir de cette trace unique nous générons 10 traces, pour chaque trace nous ne considérons que les n premières sondes envoyées à chaque saut, pour n variant de 1 à 10. Nous évaluons ensuite la qualité de la trace obtenue, en termes de sauts sans réponse. La figure 1 montre en ordonnées la fraction de chemins comportant au moins un saut sans aucune réponse parmi les n sondes envoyées (en abscisses). Intéressons-nous à la courbe pour w = 500 ms. On observe une nette amélioration de la qualité des chemins entre 1 et 3 sondes ; cela correspond à un intervalle d'une seconde, pendant lequel la plupart des problèmes temporaires sont détectés. L'amélioration est beaucoup moins nette au delà (23% de chemins incomplets avec 4 sondes, 21% avec 10 sondes). Ce plateau semble montrer que les cas temporaires ont été résolus, et qu'il ne reste plus que les cas permanents, donc les routeurs muets. Ceux-ci affectent donc une large portion des chemins, environ 20%.

Bilan : nombre de sondes et délai inter-sonde. La figure 1 (gauche) donne des indications sur le paramétrage du nombre de sondes et du délai inter-sonde. Pour un délai de 50 ms, l'amélioration est claire jusqu'à 7 sondes. Pour des délais supérieurs, 4 sondes suffisent. Nous travaillons actuellement sur l'évaluation du compromis entre le nombre de sondes, la durée et la qualité de la mesure.

Bilan : condition d'abandon. Certains outils, en particulier Paris traceroute, abandonnent la mesure d'un chemin après avoir rencontré un nombre paramétrable de sauts

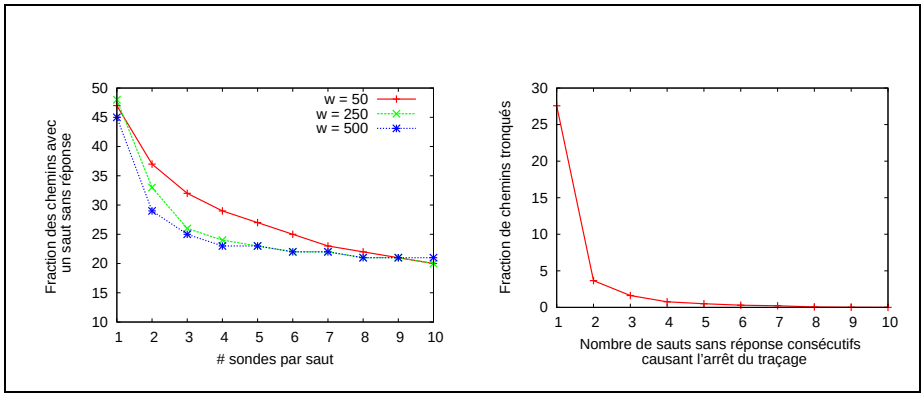

Figure 1. *A gauche : influence du nombre de sondes par saut sur la qualité du traçage, pour différents délais inter-sonde (w). A droite : influence du paramétrage de la condition d'abandon du traçage (nombre de sauts consécutifs sans réponse).*

consécutifs sans réponse, noté M. La présence de plusieurs routeurs anonymes consécutifs peut forcer l'abandon prématuré du traçage et découvrir des chemins tronqués. Pour évaluer l'impact de ce paramètre nous avons tracé les chemin vers nos 5000 destinations, en abandonnant le traçage après 10 sauts sans réponse. A partir de cette trace nous avons construit la figure 1 (droite) qui évalue le nombre de chemins tronqués (traçage abandonné trop tôt) en fonction de M. Ici encore, un compromis est nécessaire. Au delà de 3 sauts, le bénéfice n'est pas net. Notons tout de même l'existence de chemins traversant jusqu'à 9 routeurs anonymes, ce qui laisse supposer l'existence de réseaux entiers qui bloquent les sondes traceroute.

3.3. *Filtrage en périphérie*

Traceroute utilise les messages ICMP "TTL Exceeded" pour tracer la route dans le coeur du réseau. Lorsqu'il atteint sa destination, il se base sur un autre type de réponse, qui dépend du protocole utilisé : réponse d'écho ICMP, ICMP "Port Unreachable" pour UDP, SYN/ACK ou RST pour TCP. Pour des raisons de sécurité, il arrive fréquemment que l'hôte destinataire, ou encore un pare-feu à l'entrée du réseau du destinataire, bloque l'envoi de telles réponses. En conséquence, traceroute est incapable de détecter qu'il a atteint la fin du chemin, et continue à émettre des sondes avec un TTL croissant, ce qui mène à l'apparition d'une grande quantité de sauts sans réponse (sauts étoilés) en fin de chemin. Le nombre de ces sauts sans réponse en fin de chemin n'a aucune signification (contrairement aux sauts étoilés dans le coeur du réseau, ici on ne peut pas faire correspondre chaque saut à un noeud du réseau) et dépend uniquement du paramétrage de l'outil (voir section 3.2). C'est pourquoi dans l'article nous ne considérons que des traces nettoyées, c'est à dire dans lesquelles nous avons éliminé tous les sauts sans réponse en fin de chemin.

	TCP	TCP/80	UDP	ICMP
Aléatoire, répondant au ping (5000 adr.)	24	26	20	46
Aléatoire (5000 adresses)	5	5	4	6
Serveurs Web (500 adresses)	43	97	41	77
Routeurs (69000 adresses)	66	67	68	73

Tableau 2. *Fraction de destinations atteintes pour 4 types de destinations et 4 types de sondes.*

Le tableau 2 indique la fraction de destinations atteintes pour 4 types de sondes et 4 types de destinations utilisées dans des travaux précédents [AUG 06, AUG 07a]. Mis à part les serveurs Web qui répondent plutôt aux requêtes sur le port 80, il semble que les sondes ICMP produisent le plus grand nombre de réponses des destinations. UDP donne les moins bons résultats.

3.4. *Filtrage des routeurs*

Lorsqu'un filtre est installé en périphérie du réseau, il bloque l'acheminement de tout paquet interdit. Cela se traduit par l'impossibilité de tracer le chemin au delà du filtre. Cependant, il existe aussi des filtres au coeur du réseau (dans les routeurs), dont le fonctionnement est différent. Le filtre n'est pas installé au niveau du plan de routage (partie du routeur qui gère le transfert des paquets d'une interface à une autre) mais au niveau du plan de gestion. Cette partie du routeur gère le calcul des tables de routage, les messages de routage, l'interface en ligne de commande, ainsi que l'émission de messages d'erreur ICMP. Or certains routeurs appliquent un filtre sur le paquet jeté, avant de décider d'envoyer un message ICMP "TTL expiré". Cette section met en évidence ces comportements inattendus suivant deux paramètres : les numéros de port et le protocole utilisé.

3.4.1. *Numéros de port*

Le choix de la plage de ports destination utilisés par un traceroute est important car certains routeurs ne réagissent qu'à certains ports spécifiques. Pour mettre en évidence ce comportement nous lançons deux mesures simultanées, l'une utilisant le port destination 33434, l'autre le port 20000. Le but est de comparer le nombre de non-réponses obtenues en utilisant un port dans la plage du traceroute classique (33434 et suivants) et en dehors de cette plage. Nous avons répété cette mesure 10 fois, et le nombre de sauts sans réponse était systématiquement plus élevé avec le port 20000. La différence est très légère, environ 6% de non-réponses supplémentaires, ce qui montre que de tels routeurs sont très peu déployés. Nous avons isolé certains de ces routeurs (figure 2) dans nos traces. Des tests plus précis on montré qu'ils n'émettent un message d'erreur que pour les paquets dans une certaine plage de ports destination ([33434, 33534], [33400, 34400]), qui correspond à la plage utilisée par le traceroute classique.

```
13 4.68.116.145          13 4.68.116.145
14 208.50.13.193         14 * * *
15 64.209.102.154        15 64.209.102.154
```

Figure 2. *Réponse sélective du routeur au saut 14. Il répond au port 33434 (à gauche) mais reste muet au port 20000 (à droite).*

Il s'agit d'un détournement de la fonction "TTL Exceeded", initialement prévue pour informer de la perte d'un paquet à cause d'une mauvaise configuration (boucle de routage), au profit unique du traçage de route, puisque ces routeurs ne génèrent des paquets d'erreur que pour traceroute et n'informent pas des éventuels autres problèmes (ils restent muets). Nous n'avons pas pu trouver trace d'une telle fonctionnalité dans la documentation de routeurs bien connus. L'existence de plusieurs plages de ports (nous en avons trouvé deux, mais il peut en exister bien plus, dans la mesure où nous n'avons pas réalisé de tests à plus grande échelle) pourrait indiquer que la plage est configurable par l'opérateur, à moins que cela ne traduise l'existence de différents modèles de routeurs.

Bilan : numéros de ports. L'utilisation de ports destination en dehors de la plage de traceroute cache certains routeurs, ce qui peut poser problème à Paris traceroute, qui émet des sondes avec différents numéros de ports, dans le but de détecter les répartiteurs de charge [AUG 07b]. La solution consiste soit à varier les ports destination dans la plage de traceroute, soit à varier le port source, qui ne semble filtré par aucun des routeurs rencontrés.

3.4.2. *Protocole*

De même que le port destination, le choix du protocole influe sur la génération de réponses. Nous avons collecté plusieurs traces en utilisant les protocoles UDP, ICMP et TCP, pour comparer le nombre de sauts sans réponse. Contrairement à l'expérience précédente, les résultats n'ont pas permis de conclure de manière certaine. En revanche, nous avons pu montrer l'existence du phénomène en isolant manuellement trois de ces routeurs. Deux d'entre eux répondent aux sondes UDP (le protocole par défaut du traceroute classique) mais restent silencieux aux sondes ICMP et TCP. Le troisième répond aux protocoles UDP et ICMP (proposés par traceroute classique) mais pas aux sondes TCP. Ces routeurs sont donc systématiquement cachés aux utilisateurs de tcptraceroute.

Bilan : protocole. Le choix du protocole dépend du type de destination tracée. TCP mais surtout ICMP semblent donner les meilleurs résultats en termes de destinations atteintes, mais certains routeurs seront systématiquement cachés avec les outils traçage basés sur ces protocoles, comme tcptraceroute. De plus, il ne faut pas perdre de vue que l'utilisation de TCP augmente le risque de déclencher une alerte de sécurité dans un des réseaux traversés.

3.5. *Charge*

Des phénomènes temporaires autres que le rate-limiting peuvent causer des non-réponses : pertes de paquet dû au contrôle de congestion, d'autant plus que les réponses ICMP ont une faible priorité ; non-génération d'erreur ICMP dû à la surcharge temporaire d'un routeur occupé à mettre à jour sa table de routage, par exemple. Leur caractère sporadique les rend difficiles à mettre en évidence. Nous avons néanmoins observé ces phénomènes dans nos traces, par exemple certains routeurs ayant des périodes de silence à durée variable, de l'ordre de la seconde à plusieurs heures. Nous travaillons actuellement sur des techniques pour détecter ces phénomènes et les distinguer les uns des autres.

4. Travaux antérieurs

Les outils de cartographie de l'internet ont très rapidement dû faire face au problème des non-réponses et de leur impact sur la qualité des cartes obtenues. Différentes approches ont été utilisées pour palier ce problème. Certains filtrent leur jeu de données en ignorant les chemins avec un routeur muet [PAN 98, GOV 00]. L'outil utilisé par Cheswick et. al. [CHE 00] arrête le traçage d'un chemin dès qu'il rencontre un noeud qui ne répond pas. D'autres techniques sont proposées pour construire des cartes précises en présence de noeud anonymes : création d'arcs connectant les noeuds adjacents [BRO 01], heuristiques pour fusionner des noeuds anonymes en conservant la consistance de la topologie obtenue [YAO 03]. Enfin, notons que Rocketfuel [SPR 02], qui crée des cartes des systèmes autonomes, ne fait aucune mention de ce problème.

Le rate-limiting est aussi un problème documenté. Il est autorisé par les standards [BAK 95], et Govindan et. al. [GOV 02] reconnaissent son utilisation en pratique. Leur outil en tient compte en fixant un intervalle d'une seconde entre chaque sonde émise. Savage motive l'utilisation de son outil Sting [SAV 99] en montrant que les outils basés sur ICMP sont limités à cause du filtrage de ce protocole. Paris traceroute [AUG 07a] introduit un mécanisme de retransmission des sondes sans réponse pour palier ce problème.

Le problème de l'altération des réponses ICMP lors de l'envoi de sondes TCP est décrit par Malone et al. [MAL 07]. Néanmoins, le problème d'association des réponses dans les outils de traçage n'est pas évoqué.

5. Conclusion

Cet article apporte trois contributions. Premièrement, nous énumérons et classons les causes d'apparition des étoiles dans les outils de traçage. Deuxièmement, nous proposons des mesures qui permettent de mettre en évidence l'existence de chacune de

ces causes. Finalement, nous montrons comment ces résultats peuvent nous permettre de mieux paramétrer nos outils.

L'interprétation de ces résultats n'est pas triviale, et nécessite souvent un compromis entre l'effort fourni pour le traçage (en temps et en nombre de paquets émis), et la qualité des traces obtenues. Des travaux sont en cours, pour mieux évaluer ce compromis. En particulier, nous pensons qu'une caractérisation plus fine du rate-limiting nous permettrait d'apporter des améliorations majeures à notre outil, Paris traceroute, qui nécessite l'envoi massif de sondes pour découvrir les chemins en présence de répartition de charge. Nous disposons aussi d'une plate-forme constituée de routeurs hétérogènes, sur laquelle nous menons actuellement des mesures pour mieux comprendre les mécanismes de génération de messages ICMP au sein des routeurs.

6. Bibliographie

[AUG 06] AUGUSTIN B., CUVELLIER X., ORGOGOZO B., VIGER F., FRIEDMAN T., LATAPY M., MAGNIEN C., TEIXEIRA R., « Avoiding traceroute anomalies with Paris Traceroute », *Proc. ACM SIGCOMM Internet Measurement Conference, IMC*, October 2006.

[AUG 07a] AUGUSTIN B., FRIEDMAN T., TEIXEIRA R., « Measuring Load-balanced Paths in the Internet », *Proc. ACM SIGCOMM Internet Measurement Conference, IMC*, October 2007.

[AUG 07b] AUGUSTIN B., FRIEDMAN T., TEIXEIRA R., « Multipath Tracing with Paris Traceroute », *Proc. IEEE Workshop on End-to-End Monitoring, E2EMON*, May 2007.

[BAK 95] BAKER F., « Requirements for IP Version 4 Routers », RFC1812, June 1995.

[BRO 01] BROIDO A., K CLAFFY, « Internet topology : connectivity of IP graphs », *Proc. Workshop on Scalability and Traffic Control in IP Networks, SPIE ITCOM Conference*, August 2001.

[CHE 00] CHESWICK B., BURCH H., BRANIGAN S., « Mapping and visualizing the internet », *ATEC'00 : Proceedings of the Annual Technical Conference on 2000 USENIX Annual Technical Conference*, 2000.

[CHO 04] CHOI B.-Y., MOON S., ZHANG Z.-L., K. P., DIOT C., « Analysis of Point to Point Packet Delay In an Operational Network », *Proc. IEEE Infocom*, March 2004.

[GOV 00] GOVINDAN R., TANGMUNARUNKIT H., « Heuristics for Internet Map Discovery », *Proc. IEEE Infocom*, March 2000.

[GOV 02] GOVINDAN R., PAXSON V., « Estimating Router ICMP Generation Delays », *Proc. of Passive and Active Measurement Workshop, PAM*, 2002.

[MAL 07] MALONE D., LUCKY M., « Analysis of ICMP Quotations », *Proc. of Passive and Active Measurement Workshop, PAM*, 2007.

[PAN 98] PANSIOT J., GRAD D., « On routes and multicast trees in the Internet », *ACM Computer Communication Review*, January 1998.

[POS 81] POSTEL J., « Internet Control Message Protocol », RFC792, 1981.

[SAV 99] SAVAGE S., « Sting : A TCP-based Network Measurement Tool », *USENIX Symposium on Internet Technologies and Systems*, October 1999.

[SPR 02] SPRING N., MAHAJAN R., WETHERALL D., « Measuring ISP Topologies with Rocketfuel », *Proc. ACM SIGCOMM*, August 2002.

[YAO 03] YAO B., VISWANATHAN R., CHANG F., WADDINGTON D., « Topology Inference in the Presence of Anonymous Routers », *Proc. IEEE Infocom*, April 2003.