Groupe de travail Réseau
Request for Comments: 3581
Catégorie : En cours de normalisation
Traduction Claude Brière de L'Isle

J. Rosenberg, dynamicsoft
H. Schulzrinne, Columbia University

01/08/03

Extension au protocole d'initialisation de session (SIP) pour l'acheminement à réponse symétrique

Résumé
Le protocole d'initialisation de session (SIP) fonctionne sur UDP et TCP, entre autres. Lorsque utilisé avec UDP, les réponses aux demandes sont retournées à l'adresse de source d'où la demande est venue, et à l'accès écrit dans la valeur du champ d'en-tête Via supérieur de la demande. Ce comportement n'est pas désirable dans de nombreux cas, en particulier, lorsque le client est derrière un traducteur d'adresse réseau (NAT, *Network Address Translator*). La présente extension définit un nouveau paramètre pour le champ d'en-tête Via, appelé "rport", qui permet à un client de demander au serveur de renvoyer la réponse à l'adresse IP de source et l'accès duquel la demande a été générée.

1. Introduction

Le protocole d'initialisation de session (SIP) [RFC3261] fonctionne sur UDP et TCP. Lorsque il est utilisé avec UDP, les réponses aux demandes sont retournées à l'adresse de source d'où vient la demande, et à l'accès écrit dans la valeur du champ d'en-tête Via supérieur de la demande. Il en résulte une façon "hybride" de calculer la destination de la réponse. La moitié des informations (précisément, l'adresse IP) est tirée des en-têtes de paquet IP, et l'autre moitié (l'accès) des en-têtes de messages SIP. SIP fonctionne de cette manière pour qu'un serveur puisse écouter tous les messages, à la fois demandes et réponses, sur une seule adresse et accès IP. Cela aide à améliorer l'adaptabilité. Cependant, ce comportement n'est pas désirable dans de nombreux cas, en particulier lorsque le client est derrière un NAT. Dans ce cas, la réponse ne va pas correctement traverser le NAT, car elle ne va pas correspondre au lien établi avec la demande.

De plus, il n'y a actuellement aucun moyen pour un client d'examiner une réponse et déterminer l'accès de source qu'a vu le serveur dans la demande correspondante. Actuellement, SIP fournit au client l'adresse IP de source que le serveur a vu dans la demande, mais pas l'accès. L'adresse IP de source est portée dans le paramètre "received" dans la valeur du champ d'en-tête Via supérieur de la réponse. Cette information s'est révéle utile pour la traversée de NAT de base, les besoins de débogage, et la prise en charge des hôtes multi rattachement. Cependant, c'est incomplet sans les informations d'accès.

La présente extension définit un nouveau paramètre pour le champ d'en-tête Via, appelé "rport", qui permet à un client de demander au serveur de renvoyer la réponse à l'adresse et accès IP de source d'où est venue la demande. Le paramètre "rport" est analogue au paramètre "received", sauf que "rport" contient un numéro d'accès, et non l'adresse IP.

2. Terminologie

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans la [RFC2119] et indiquent les niveaux d'exigence pour les mises en œuvre conformes.

3. Comportement du client

Le comportement de client spécifié ici affecte le traitement du transport défini au paragraphe 18.1 de SIP [RFC3261].

Un client, conforme à la présente spécification (les clients incluent les UAC et les mandataires) PEUT inclure un paramètre "rport" dans la valeur de champ d'en-tête Via supérieure des demandes qu'il génère. Ce paramètre DOIT n'avoir pas de valeur ; il sert de fanion pour indiquer au serveur que cette extension est prise en charge et demandée pour la transaction.

Lorsque le client envoie la demande, si la demande est envoyée en utilisant UDP, le client DOIT être prêt à recevoir la réponse sur les mêmes adresses IP et accès qu'il a utilisé pour remplir l'adresse IP de source et l'accès de source de la demande. Pour la rétro compatibilité, le client DOIT quand même être prêt à recevoir une réponse sur l'accès indiqué dans le champ "sent-by" de la valeur du champ d'en-tête Via supérieur, comme spécifié au paragraphe 18.1.1 de SIP [RFC3261].

Lorsque il y a un NAT entre le client et le serveur, la demande va créer (ou rafraîchir) un lien dans le NAT. Ce lien doit rester en vigueur pour la durée de la transaction afin que le client reçoive la réponse. La plupart des liens de NAT UDP paraissent avoir une temporisation d'environ une minute. Cela excède la duraée des transactions non INVITE. Donc, les réponses à une demande non INVITE seront reçues alors que le lien existe toujours. Les transactions INVITE peuvent prendre un temps arbitrairement long pour se terminer. Il en résulte que le lien peut être arrivé à expiration avant qu'une réponse finale ait été reçue. Pour garder la fraîcheur du lien, le client DEVRAIT retransmettre son INVITE à peu près toutes les 20 secondes. Ces retransmissions devront avoir lieu même après la réception d'une réponse provisoire.

Un UA PEUT exécuter l'algorithme de découverte de la durée de vie de lien du paragraphe 10.2 de la [RFC3489] pour déterminer la durée de vie réelle du lien dans le NAT. Si elle est supérieure à une minute, le client DEVRAIT augmenter l'intervalle des retransmissions de demande jusqu'à la moitié de la durée de vie découverte. Si elle est de moins d'une minute, il DEVRAIT diminuer l'intervalle des retransmissions de demande à la moitié de la durée de vie découverte. Noter que cette découverte des durées de vie de liens peut n'être pas fiable. Voir le paragraphe 14.3 de la [RFC3489].

4. Comportement du serveur

Le comportement du serveur spécifié ici affecte le traitement du transport défini au paragraphe 18.2 de SIP [RFC3261].

Lorsque un serveur conforme à la présente spécification (qui peut être un mandataire ou un UAS) reçoit une demande, il examine la valeur du champ d'en-tête Via supérieur. Si cette valeur de champ d'en-tête Via contient un paramètre "rport" sans aucune valeur, il DOIT régler la valeur du paramètre à l'accès de source de la demande. Ceci est analogue à la façon dont un serveur va insérer le paramètre "received" dans la valeur du champ d'en-tête Via supérieur. En fait, le serveur DOIT insérer un paramètre "received" contenant l'adresse IP de source d'où venait la demande, même si elle est identique à la valeur du composant "sent-by". Noter que ce traitement a lieu indépendamment du protocole de transport.

Lorsque un serveur tente d'envoyer une réponse, il examine la valeur du champ d'en-tête Via supérieur de cette réponse. Si le composant "sent-protocol" indique un protocole de transport d'envoi individuel non fiable, comme UDP, et si il n'y a pas de paramètre "maddr", mais si il y a à la fois un paramètre "received" et un paramètre "rport", la réponse DOIT être envoyée à l'adresse IP mentionnée dans le paramètre "received", et à l'accès mentionné dans le paramètre "rport". La réponse DOIT être envoyée de la même adresse et accès que la demande correspondante reçue. Cela ajoute effectivement une nouvelle étape de traitement entre les étapes deux et trois du paragraphe 18.2.2 de SIP [RFC3261].

La réponse doit être envoyée de la même adresse et accès de réception que la demande afin de traverser les NAT symétriques. Lorsque un serveur écoute les demandes sur plusieurs accès ou interfaces, il aura besoin de se souvenir de celui sur lequel la demande a été reçue. Pour un mandataire à états pleins, la mémorisation de ces informations pour la durée de la transaction n'est pas un problème. Cependant, un mandataire sans états ne mémorise pas l'état entre une demande et sa réponse, et ne peut donc pas se souvenir de l'adresse et de l'accès sur lesquels une demande a été reçue. Pour mettre correctement en œuvre la présente spécification, un mandataire sans état peut coder l'adresse et accès de destination d'une demande dans la valeur du champ d'en-tête Via qu'il insère. Quand la réponse arrive, il peut extraire ces informations et les utiliser pour transmettre la réponse.

5. Syntaxe

La syntaxe du paramètre "rport" est :

response-port = "rport" [EQUAL 1*DIGIT]

Cela étend la définition existante des paramètres du champ d'en-tête Via, de sorte que son BNF est maintenant :

via-params = via-ttl / via-maddr / via-received / via-branch / response-port / via-extension

6. Exemple

Un client envoie à un serveur mandataire un INVITE qui ressemble, en partie, à ceci :

INVITE sip:user@example.com SIP/2.0
Via: SIP/2.0/UDP 10.1.1.1:4540;rport;branch=z9hG4bKkjshdyff

Cet INVITE est envoyé avec un accès de source de 4540 et une adresse IP de source de 10.1.1.1. Le mandataire est à 192.0.2.2 (proxy.example.com), il écoute sur les deux accès 5060 et 5070. Le client envoie la demande à l'accès 5060. La demande passe à travers un NAT sur le chemin vers le mandataire, de sorte que l'adresse IP de source apparaît comme 192.0.2.1 et l'accès de source comme 9988. Le mandataire transmet la demande, mais pas avant d'ajouter une valeur au paramètre "rport" dans la demande relayée :

INVITE sip:user@example.com SIP/2.0
Via: SIP/2.0/UDP proxy.example.com;branch=z9hG4bKkjsh77
Via: SIP/2.0/UDP 10.1.1.1:4540;received=192.0.2.1;rport=9988;branch=z9hG4bKkjshdyff

Cette demande génère une réponse qui arrive au mandataire :

SIP/2.0 200 OK
Via: SIP/2.0/UDP proxy.example.com;branch=z9hG4bKkjsh77
Via: SIP/2.0/UDP 10.1.1.1:4540;received=192.0.2.1;rport=9988;branch=z9hG4bKkjshdyff

Le mandataire supprime sa valeur de champ d'en-tête Via supérieur, et examine alors la suivante. Elle contient à la fois un paramètre "received" et un paramètre "rport". Le serveur suit les règles spécifiées à la Section 4 et envoie la réponse à l'adresse IP 192.0.2.1, accès 9988, et l'envoie de l'accès 5060 sur 192.0.2.2:

SIP/2.0 200 OK
Via: SIP/2.0/UDP 10.1.1.1:4540;received=192.0.2.1;rport=9988;branch=z9hG4bKkjshdyff

Ce paquet satisfait au lien créé par la demande initiale. Donc, le NAT réécrit l'adresse de destination de ce paquet comme 10.1.1.1, et l'accès de destination comme 4540. Il transmet cette réponse au client, qui est en train d'écoute la réponse sur cette adresse et cet accès. Le client reçoit correctement la réponse.

7. Considérations sur la sécurité

Lorsque un serveur utilise la présente spécification, les réponses qu'il envoie vont maintenant inclure l'accès de source d'où est venue la demande. Dans certaines instances, l'adresse et l'accès de source d'une demande sont des informations sensibles. Si elles le sont, les demandes DEVRAIENT être protégées en utilisant SIP sur TLS [RFC3261]. Dans un tel cas, la présente spécification ne fournit aucune fonction d'acheminement de la réponse (parce que cela ne fonctionne qu'avec TCP) ; cela donne seulement au client des informations sur l'accès de source tel que vu par le serveur.

Il est possible qu'un attaquant essaye d'interrompre le service à un client en agissant par interposition, en modifiant le paramètre "rport" dans l'en-tête Via de la demande envoyée par un client. La suppression de ce paramètre va empêcher les clients derrière des NAT de recevoir le service. L'ajout du paramètre n'aura généralement pas d'impact. Bien sûr, si un attaquant est capable de lancer une attaque par interposition, il a beaucoup d'autres moyens de faire du déni de service, comme de simplement supprimer la demande. Cette attaque ne semble donc pas très significative.

8. Considérations relatives à l'IANA

Aucune considération relative à l'IANA n'est associée à la présente spécification.

9. Considérations relatives à l'IAB

L'IAB a étudié une classe de protocoles auxquels on se réfère sous le nom de protocoles d'auto réparation d'adressage unilatéral (UNSAF, *Unilateral Self Address Fixing*) [RFC3424]. Ces protocoles permettent à un client derrière un NAT d'apprendre l'adresse et l'accès IP que le NAT va allouer pour une certaine demande, afin d'utiliser cette information dans des protocoles de couche application. Un exemple de protocole UNSAF est la simple traversée par le protocole de datagramme d'utilisateur (UDP) des traducteurs d'adresse réseau (STUN, *Simple Traversal of UDP Through NATs*) [RFC3489].

Tout protocole est un protocole UNSAF si il révèle à un client, l'adresse et l'accès IP de source d'un paquet envoyé à travers ce NAT. Bien qu'elle ne soit pas conçue pour ce propos, la présente spécification peut être utilisée comme un protocole UNSAF. En utilisant le paramètre "rport" (défini ici) et le paramètre "received" (défini dans la [RFC3261]) dans la valeur du champ d'en-tête Via supérieur d'une réponse, un client qui envoie une demande peut apprendre l'adresse comme elle a été vue par le serveur qui a envoyé la réponse.

Cette information a deux utilisations. La première est pour les enregistrements. Considérons un client derrière un NAT qui souhaite s'enregistrer auprès d'un mandataire/registraire de l'autre côté du NAT. Le client doit fournir, dans son enregistrement, l'adresse à laquelle il devrait recevoir les demandes SIP entrantes de la part du mandataire. Cependant, comme le client est situé derrière un NAT, aucune des adresses ne sera accessible sur aucune de ses interfaces depuis le mandataire. Si le client peut fournir au mandataire une adresse que le mandataire puisse atteindre, le client pourra recevoir les demandes entrantes. En utilisant la présente spécification, un client derrière un NAT peut apprendre son adresse et son accès tels que vus par le mandataire qui reçoit une demande REGISTER. Le client peut alors effectuer un enregistrement supplémentaire, en utilisant cette adresse dans un en-tête Contact. Cela permettrait à un client de recevoir des demandes entrantes, comme une INVITE, sur l'adresse et l'accès IP qu'il a utilisé pour remplir l'adresse et l'accès IP de source de la demande d'enregistrement qu'il a envoyé. Cette approche ne va fonctionner que lorsque les serveurs envoient des demandes à un UA à partir de la même adresse et accès que ceux du REGISTER qui a lui-même été reçu.

Dans de nombreux cas, le serveur auquel l'enregistrement est envoyé ne sera pas le registraire lui-même, mais plutôt un mandataire qui envoie ensuite la demande au registraire. Dans un tel cas, toute demande entrante pour le client doit traverser le mandataire auquel l'enregistrement a été directement envoyé. L'extension d'en-tête Path à SIP [RFC3327] permet au mandataire d'indiquer qu'il doit être sur le chemin de telles demandes.

Le second usage est pour l'acheminement d'enregistrement, pour régler le même problème que ci-dessus, mais entre deux mandataires. Un mandataire derrière un NAT qui transmet une demande à un serveur peut utiliser OPTIONS, par exemple, pour apprendre son adresse telle que vue par ce serveur. Cette adresse peut être placée dans le champ d'en-tête Record-Route des demandes envoyées à ce serveur. Cela va permettre au mandataire de recevoir les demandes provenant de ce serveur sur la même adresse et accès IP qu'utilisés pour remplir l'adresse et accès IP de source de la demande OPTIONS.

À cause de cet usage potentiel, le présent document doit examiner les problèmes soulevés dans la [RFC3424].

9.1 Définition du problème

D'après la [RFC3424], toute proposition d'UNSAF doit fournir :

Une définition précise d'un problème spécifique, de portée limitée, qui a été résolu par la proposition d'UNSAF. Une solution à court terme ne devrait pas être généralisée pour résoudre d'autres problèmes ; c'est pourquoi les "solutions à court terme ne le sont généralement pas".

La présente spécification est principalement destinée à permettre aux réponses SIP d'être reçues lorsque une demande est envoyée à travers un NAT. Dans cette application principale, la présente specification n'est pas une proposition d'UNSAF. Cependant, par un effet collatéral de cette capacité, la présente spécification peut être utilisée comme un protocole UNSAF. Dans cet usage, elle va régler deux problèmes :

- Fournir à un client une adresse qu'il peut utiliser dans l'en-tête Contact d'une demande REGISTER quand il est derrière un NAT.
- Fournir à un mandataire une adresse qu'il peut utiliser dans un en-tête Record-Route d'une demande, quand il est derrière un NAT.

9.2 Stratégie de sortie

D'après la [RFC3424], toute proposition d'UNSAF doit fournir :

La description d'une stratégie de sortie/plan de transition. Les meilleures solutions à court terme sont celles qui vont naturellement avoir de moins en moins d'utilité à mesure que la technologie appropriée est déployée.

Le groupe de travail SIP a reconnu que l'usage de la présente spécification n'est pas approprié pour la prise en charge des enregistrements et acheminements d'enregistrement à travers des NAT. Elle pose un certain nombre de problèmes connus qui sont documentés ci-dessous. La façon d'éliminer l'usage potentiel de la présente spécification pour la réparation d'adresse est de fournir une solution appropriée aux problèmes qui pourraient motiver l'usage de cette spécification pour la réparation d'adresse. Précisément, les solutions appropriées pour les enregistrements et l'acheminement d'enregistrement en présence de NAT doivent être développées. Ces solutions ne vont pas s'appuyer sur la réparation d'adresse.

Les exigences pour de telles solutions sont déjà en cours de développement [RFC5923].

Les mises en œuvre de la présente spécification sont encouragées à suivre ce travail pour de meilleures solutions pour les enregistrements et l'acheminement d'enregistrement à travers un NAT.

9.3 Fragilité introduite par la présente spécification

D'après la [RFC3424], toute proposition d'UNSAF doit fournir :

La discussion de questions spécifiques qui peuvent rendre les systèmes plus "fragiles". Par exemple, les approches qui impliquent l'utilisation de données à plusieurs couches réseau créent plus de dépendance, augmentent les difficultés de débogage, et rendent plus difficile la transition.

La présente spécification, si elle est utilisée pour la réparation d'adresse, introduit plusieurs points de fragilité dans un système SIP :

- Si il est utilisé pour des enregistrements UDP, un client aura besoin de se réenregistrer fréquemment afin de conserver la fraîcheur des liens de NAT. Dans de nombreux cas, ces enregistrements vont devoir avoir lieu presque cent fois plus souvent que l'intervalle de rafraîchissement normal d'un enregistrement. Cela introduit une charge sur le système et altère l'adaptabilité.

- Un client ne peut pas déterminer précisément la durée de vie du lien d'un NAT au travers duquel il s'enregistre (ou achemine l'enregistrement). Donc, il y a des périodes d'inaccessibilité qui se produisent entre le moment où un lien arrive à expiration et celui de l'envoi du prochain enregistrement ou rafraîchissement d'OPTIONS. Il peut en résulter des appels, messages, ou autres informations, manqués.

- Si le NAT est de la variété symétrique [RFC3489], un client sera seulement capable d'utiliser son adresse pour recevoir des demandes provenant du serveur auquel il a envoyé la demande. Si ce serveur fait partie d'une grappe de nombreux serveurs, le client peut n'être pas capable de recevoir les demandes provenant d'autres serveurs de la grappe. Il peut en résulter des appels, messages, ou autres informations, manqués.

- Si le NAT est de la variété symétrique [RFC3489], un client sera seulement capable d'utiliser son adresse pour recevoir des demandes si le serveur envoie les demandes au client à partir de la même adresse et accès d'où le serveur a reçu les enregistrements. Ce comportement de serveur n'est pas rendu obligatoire par la [RFC3261], bien qi'l apparaisse de façon courante en pratique.

- Si le registraire et le serveur auxquels le client a envoyé sa demande REGISTER ne sont pas les mêmes, l'approche ne va fonctionner que si le serveur utilise le champ d'en-tête Path [RFC3327]. Il n'y a pas de façon aisée et fiable pour que le serveur détermine que l'en-tête Path devrait être utilisé pour un enregistrement. Utiliser Path quand l'adresse dans le champ d'en-tête Via supérieur est une adresse privée va généralement fonctionner, mais peut résulter en une utilisation de Path alors qu'elle n'est pas réellement nécessaire.

9.4 Exigences pour une solution à long terme

D'après la [RFC3424], toute proposition d'UNSAF doit fournir :

L'identification des exigences pour des solutions à plus long terme, techniquement viables -- contribue au processus de découverte de la bonne solution à plus long terme.

La fragilité décrit au paragraphe 9.3 nous a conduit aux exigences suivantes pour une solution à long terme :

Le client ne devrait pas avoir à spécifier son adresse. Les enregistrements et acheminements d'enregistrement exigent que le client spécifie l'adresse à laquelle il devrait recevoir les demandes. Une solution technique viable devrait permettre au client de spécifier explicitement qu'il veut recevoir les demandes entrantes sur la connexion sur laquelle a été envoyée la demande sortaante. De cette façon, le client n'a pas besoin de spécifier son adresse.

La solution doit prendre en compte les grappes de serveurs. Dans de nombreux systèmes SIP commercialisés, il y aura de nombreux serveurs, chacun à des adresses et accès différents, qui traitent les demandes entrntes pour un client. La solution doit explicitement considérer ce cas.

La solution ne doit pas exiger une augmentation de la charge du réseau. Il ne peut pas y avoir une pénalisation associée à une bonne solution technique.

9.5 Problèmes avec les boîtes NAPT existantes

D'après la [RFC3424], toute proposition d'UNSAF doit fournir :

La discussion de l'impact des problèmes pratiques notés avec les NA[P]T existants, déployés et les rapports d'expériences.

À notre connaissance, au moment de cette rédaction, il y a seulement un nombre très limité d'usages de la présente spécification pour la réparation d'adresse. Donc, aucun problème pratique spécifique n'a été soulevé.

10. Remerciements

Les auteurs tiennent à remercier Rohan Mahy et Allison Mankin de leurs commentaires et contributions à ce travail.

11. Références

11.1 Références normatives

[RFC2119] S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997.

[RFC3261] J. Rosenberg et autres, "SIP : Protocole d'initialisation de session", juin 2002. (*Mise à jour par* RFC3265, RFC3853, RFC4320, RFC4916, RFC5393, RFC6665)

[RFC3327] D. Willis, B. Hoeneisen, "Champ d'en-tête d'extension du protocole d'initialisation de session (SIP) pour enregistrer des contacts non adjacents", décembre 2002. *(P.S.)*

[RFC3489] J. Rosenberg et autres, "STUN - Simple traversée par le protocole de datagramme d'utilisateur (UDP) des traducteurs d'adresse réseau (NAT)", mars 2003. (*Obsolète, voir* RFC5389) *(P.S.)*

11.2 Références pour information

[RFC3424] L. Daigle, éd., IAB, "Considérations de l'IAB sur la réparation d'auto adressage unilatéral (UNSAF) à travers la traduction d'adresse réseau", novembre 2002. *(Information)*

[RFC5923] V. Gurbani, R. Mahy, etc., "Réutilisation de connexion dans le protocole d'initialisation de session (SIP)", juin 2010. *(P.S.)*

12. Déclaration de propriété intellectuelle

ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur le répertoire en ligne des IPR de l'IETF à http://www.ietf.org/ipr.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à ietf-ipr@ietf.org.

13. Adresse des auteurs

Jonathan Rosenberg
dynamicsoft
600 Lanidex Plaza
Parsippany, NJ 07054
USA
téléphone : +1 973 952-5000
mél : jdrosen@dynamicsoft.com
URI : http://www.jdrosen.net

Henning Schulzrinne
Columbia University
M/S 0401
1214 Amsterdam Ave.
New York, NY 10027
USA
mél : schulzrinne@cs.columbia.edu
URI : http://www.cs.columbia.edu/~hgs

14. Déclaration complète de droits de reproduction

Remerciement
Le financement de la fonction d'édition des RFC est actuellement fourni par l'Internet Society.