

RFC 6076 : Basic Telephony SIP End-to-End Performance Metrics

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 20 janvier 2011

Date de publication du RFC : Janvier 2011

<https://www.bortzmeyer.org/6076.html>

Le protocole de signalisation SIP est aujourd'hui le plus utilisé pour la voix sur IP. Il existe de nombreux fournisseurs SIP concurrents, de nombreux services SIP un peu partout. Comment les comparer ? Ce RFC se consacre à définir des **métriques** pour la téléphonie SIP.

Il est le premier RFC produit par le groupe de travail PMOL <<http://tools.ietf.org/wg/pm01>> de l'IETF, groupe de travail qui est consacré à la définition de métriques pour les couches hautes (notamment la couche 7), le groupe IPPM <<http://tools.ietf.org/wg/ippm>> se consacrant aux couches basses comme la couche 3. En suivant ses recommandations, il devrait être possible de comparer des mesures de performances SIP entre différents services. À noter qu'il existait déjà des métriques pour des systèmes de téléphonie autres, comme SS7, mais rien pour SIP.

Comme PMOL (contrairement à un autre groupe de travail sur les mesures, BMWG <<http://tools.ietf.org/wg/bmwg>>) établit des métriques pour le monde extérieur, le grand Internet, les résultats des mesures sont influencés par des variables extérieures à SIP, par exemple la qualité du réseau (latence et capacité libre). Les métriques définies ici n'essaient pas de séparer ces variables « réseau » des autres. On teste donc un service, pas uniquement un serveur.

D'autre part, il s'agit bien de définir des métriques (des grandeurs mesurables et définies rigoureusement) et pas des objectifs chiffrés, genre SLA. Ce RFC définit ainsi ce qu'est exactement un temps de réponse, mais pas quelle valeur de ce temps est acceptable pour l'utilisateur.

Après ces préliminaires, la section 2 définit le vocabulaire utilisé. Ainsi, "*End-to-end*" fait référence à une liaison entre les deux entités qui communiquent (deux téléphones, par exemple), même si cette communication passe, au niveau 3 par beaucoup de routeurs, et au niveau 7 par plusieurs relais SIP. Une **session** est la communication établie par SIP (dont le rôle principal est justement d'établir et de fermer

les sessions; le transfert de la voix encodée n'est pas directement du ressort de SIP) et elle peut donc comprendre plusieurs flots de données différents, par exemple plusieurs connexions TCP.

Enfin (non, tout le travail préparatoire n'est pas fini, on n'est pas encore arrivé aux métriques elles-mêmes), toute mesure faisant intervenir le temps, la section 3 est dédiée aux problèmes de mesure de celui-ci. C'est essentiellement une reprise du RFC 2330¹, section 10. Pour les besoins des mesures SIP (moins exigeantes que les mesures comme le temps d'acheminement d'un paquet IP du RFC 7679), une erreur de moins d'une milli-seconde est demandée, ainsi qu'une stabilité de un pour mille pour la fréquence de l'horloge.

La section 4 introduit enfin ces fameuses métriques. Le temps commence lorsque le premier message est envoyé par un des UA ("*User Agents*", typiquement un téléphone, matériel ou logiciel). Même si ce premier message doit être répété, l'horloge n'est pas remise à zéro. On arrête le chrono lorsque l'opération souhaitée par l'UA est faite, même si elle a nécessité plusieurs transactions SIP (par exemple, si un appel a été redirigé). Des options comme l'authentification (qui est fréquemment utilisée dans le monde SIP car les fournisseurs de service qui procurent un accès au téléphone classique ne sont jamais gratuits) peuvent sérieusement augmenter la durée mesurée mais on n'essaie pas de les retirer du total : elles font partie du service mesuré (voir section 5.3 pour les détails). Enfin, les métriques de la section 4 ne précisent pas la taille de l'échantillon utilisé, se contentant de rappeler qu'on a en général de meilleurs résultats en augmentant la taille de l'échantillon.

Pour comprendre les métriques, il faut se rappeler que le fonctionnement typique d'un téléphone SIP est :

- Au démarrage du téléphone, on s'enregistre auprès d'un fournisseur SIP (SSP pour "*SIP Service Provider*") tel que OVH ou Free/Freephonie. Cela permettra, notamment, de recevoir des appels entrants même si l'appelant ne connaît pas notre adresse IP actuelle.
- Pour chaque appel téléphonique, on établit une session SIP avec le destinataire, et cette session sert à transmettre les paramètres d'envoi du flux audio (qui n'est pas envoyé dans la session SIP mais via un protocole séparé comme RTP).

Parmi les métriques définies, on trouve :

- RRD ("*Registration Request Delay*", section 4.1). Elle mesure le temps écoulé entre un enregistrement d'un UA auprès d'un fournisseur SIP (requête REGISTER) et la réponse réussie, typiquement 200 OK (les échecs sont comptés différemment, voir section 4.2), exprimée en milli-secondes.
- IRA ("*Ineffective Registration Attempts*", section 4.2) qui est le pourcentage des tentatives d'enregistrement qui échouent, soit par réponse négative, soit par expiration du délai d'attente.
- SRD ("*Session Request Delay*", section 4.3), qui est le délai nécessaire pour établir une nouvelle session. En termes de téléphonie classique, c'est le moment entre la composition du numéro et le décrochage du téléphone en face. En SIP, c'est le délai entre la requête INVITE et la réponse à celle-ci (la réponse pouvant être négative, par exemple un 486 - Occupé).
- SDT ("*Session Duration Time*", section 4.5) est la durée d'une session (d'un appel). C'est la différence entre le moment de la fin de la session (par requête explicite ou bien par expiration du délai d'attente) et la réponse réussie à un INVITE.
- SER ("*Session Establishment Ratio*", section 4.6) est le pourcentage de tentatives d'établissement de sessions SIP qui réussit (de même que IRA était le pourcentage de tentatives d'enregistrement auprès du SSP qui échouaient).

Le logiciel de test sipsak <<http://sipsak.org/>> affiche dans certains cas des durées mais elles ne sont pas exprimées en terme des métriques de ce RFC donc il faudrait lire le source pour avoir leur sémantique exacte. Ici, l'établissement d'une session :

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc2330.txt>

```
% sipsak -vv -s "sip:bortzmeyer@ekiga.net"
...
message received:
SIP/2.0 200 OK
...
** reply received after 6.770 ms **
```

Et ici une session complète avec envoi d'un message :

```
% sipsak -vv -B "This is a test" -M -s "sip:stephane@10.10.86.76"
...
received last message 55.642 ms after first request (test duration).
```