

RFC 3920 : Extensible Messaging and Presence Protocol (XMPP): Core

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 avril 2008

Date de publication du RFC : Octobre 2004

<https://www.bortzmeyer.org/3920.html>

Jabber est le nom original, et XMPP le nom IETF d'un protocole d'échange d'informations encodées en XML, protocole surtout connu pour son utilisation dans la messagerie instantanée.

La messagerie instantanée est une application très populaire surtout chez les moins de dix ans. Elle permet d'échanger des messages sans prendre le temps de réfléchir et en général sans craindre qu'ils soient archivés (contrairement au courrier électronique où il faut réfléchir car ce que l'on écrit restera). Elle sert à de nombreux usages, discussions entre "*geeks*", systèmes de surveillance avec notification des problèmes, etc. À l'exception du vénérable protocole IRC (décrit - insuffisamment - dans le RFC 2810¹ et suivants), les systèmes existants de messagerie instantanée sont tous basés sur des protocoles privés et fermés comme MSN. La normalisation qui fait le succès du courrier électronique n'a pas réussi ici.

D'où le vieux projet (RFC 2779) de développer un protocole moderne et ouvert, adapté à cette application. C'est ce qu'on fait les développeurs du protocole Jabber en 1999, projet adopté ultérieurement par l'IETF, rebaptisé XMPP en 2002 et qui était normalisé dans ce RFC (l'annexe D liste les différences entre Jabber et XMPP, la plupart liées à la sécurité et à l'internationalisation). Depuis, il a été remplacé par le RFC 6120. Notre RFC décrit un protocole généraliste, un RFC compagnon, le RFC 3921, normalise les aspects spécifiques à la messagerie instantanée.

Quels sont les principes de XMPP? Les données sont encodées en XML (avec quelques bémols notés en section 11), dans des **strophes** ("*stanzas*") qui sont des petits paquets d'information. Ces strophes sont transmises sur une connexion TCP (section 4.2) entre client et serveur, ou bien entre serveur et serveur. Les entités qui communiquent (ce ne sont pas forcément des humains, Jabber ayant été conçu pour permettre également la communication entre programmes) sont identifiées par une adresse, qui

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc2810.txt>

a une forme syntaxique proche de celle du courrier électronique. Par exemple, mon adresse XMPP est `bortzmeyer@gmail.com`.

XMPP, comme le courrier électronique, fonctionne par le connexion de serveurs, chacun chargé d'un ou plusieurs domaines. Les clients transmettent les messages aux serveurs, qui se chargent de les acheminer au serveur suivant (section 2.1 du RFC).

Les adresses (JID pour "*Jabber Identifieur*") sont décrites dans la section 3. Outre le classique `nom@domaine`, elles peuvent comporter une **ressource**, indiquée après la barre oblique (par exemple `bortzmeyer@gmail.com/AFN`).

La représentation des données en XML est décrite dans la section 4. Un flux de données, un "*stream*", est un **élément** XML `<stream>`, qui comporte une ou plusieurs strophes, `<stanza>`. Il existe de nombreux types de strophes (section 9) sans compter les messages d'erreur, décrits en 4.7. Voici un exemple classique d'un échange (C étant le client et S le serveur), avec des strophes de type `message` :

```
C: <message from='juliet@example.com'
      to='romeo@example.net'
      xml:lang='en'>
C:   <body>Art thou not Romeo, and a Montague?</body>
C: </message>
S: <message from='romeo@example.net'
      to='juliet@example.com'
      xml:lang='en'>
S:   <body>Neither, fair saint, if either thee dislike.</body>
S: </message>
```

Contrairement à beaucoup d'autres protocoles de messagerie instantanée, conçus uniquement pour jouer, XMPP, étant prévu aussi pour le travail, dispose de fonctions de sécurité très riches. La section 5 précise l'utilisation de TLS et la 6 celle de SASL.

La section 9 détaille les types de strophes disponibles. Trois types sont définis dans ce RFC, `<message>`, `<presence>` et `<iq>`. `<message>`, dont la sémantique est décrite en section 9.2.1, transporte un... message, tandis que `<iq>` (section 9.2.3) est une interrogation ("*Info / Query*"). Ces éléments XML ont un certain nombre d'**attributs** par exemple l'attribut `to`, section 9.1.1, qui spécifie l'adresse du destinataire (`romeo@example.net` dans la première strophe de l'exemple ci-dessus). La syntaxe complète, exprimée en W3C Schema, figure dans l'annexe C.

XMPP est aujourd'hui mis en œuvre dans de nombreux logiciels. C'est ainsi que le service de messagerie instantanée Google Talk utilise XMPP. Le client XMPP (qui est aussi un client de nombreux autres protocoles) le plus populaire sur Unix est sans doute Pidgin (ex-Gaim). Côté serveurs, le plus pittoresque (et pas le moins utilisé) est sans doute ejabberd, écrit en Erlang.

Les développeurs liront avec intérêt le livre "*Programming Jabber*" <https://www.bortzmeyer.org/programming-jabber.html> qui explique comment utiliser Jabber/XMPP dans ses propres programmes.