

1. Présentation théorique

1. Introduction.

Le **protocole SMTP** (*Simple Mail Transfer Protocol*, traduisez *Protocole Simple de Transfert de Courrier RFC 821*) est le protocole standard permettant de transférer le courrier d'un serveur à un autre en connexion point à point. Il s'agit d'un protocole fonctionnant en mode connecté, encapsulé dans une trame TCP/IP. Le courrier est remis directement au serveur de courrier du destinataire. Le protocole SMTP fonctionne grâce à des commandes textuelles envoyées au serveur SMTP (par défaut sur le port 25). Chacune des commandes envoyées par le client (validée par la chaîne de caractères *CR/LF*, équivalent à un appui sur la touche entrée) est suivi d'une réponse du serveur SMTP composée d'un numéro et d'un message descriptif.

Prenons comme exemple l'utilisateur `user1@site1.be` désirant envoyer un courrier à l'utilisateur `user2@site2.be`. L'utilisateur "user1" a configuré son logiciel de messagerie en y renseignant le serveur de courrier avec lequel il désire travailler: soit le serveur `smtp.site1.be`. Tout courrier sortant sera envoyé vers ce serveur qui a son tour devra prendre contact dans notre exemple avec le serveur SMTP enregistré dans le serveur DNS du domaine `site2.be` sous le record MX. Considérons maintenant l'échange qui va se produire entre les deux serveurs:

Lors de la connexion sur le serveur `smtp.site2.be` sur le port 25 en mode TCP, celui-ci envoie le code 220 suivi d'une chaîne de caractères optionnelle libre indiquant qu'il est prêt à accepter la réception du courrier

```
220 smtp.site2.be SMTP sendmail ready
helo smtp.site1.be (ou alors site1.be)
250 hello smtp.site2.be
MAIL FROM:user1@site1.be
250 ok
RCPT TO: user2@site2.be
250 ok its for <user2@site2.be>
DATA
354 ok, send it; end with <CRLF>.<CRLF>
from: user1@site1.be
to: user2@site2.be
date: 17/11/04
subject: titre du message
<CRLF>
Ceci est le corps du message que je dois envoyer
<CRLF>.<CRLF>
250 ok
QUIT
221smtp.site2.be closing transmission
```

Les spécifications de base du protocole SMTP veulent que tous les caractères transmis soient codés sur 7 bits et que le 8^{ème} bit soit explicitement mis à zéro. Ainsi pour envoyer des caractères accentués il faut faire recours à des algorithmes intégrant les spécifications MIME:

- **base64** pour les fichiers attachés
- **quoted-printable** (d'abréviation *QP*) pour les caractères spéciaux contenus dans le corps du message

Pour éviter de devoir utiliser ces techniques pour l'utilisation des caractères accentués, une extension de cette norme trop simple a été proposée sous la norme RFC 1425. Cette extension entrevoit la possibilité d'ajout d'extensions dont en voici certaines:

- la norme RFC 1427 qui permet au serveur de destination de pouvoir renseigner la taille limite des courriers qu'il peut accepter. Avant cette norme, la seule possibilité pour un serveur est d'accepter le message et ensuite de le détruire.

- la norme RFC 1426 qui permet de pouvoir envoyer les corps du message en utilisant un codage ASCII sur 8 bits, acceptant alors les caractères accentués.

Pour pouvoir se connecter sur un serveur de courriers en mode SMTP où ESMTP, il suffit d'utiliser la commande texte HELO pour le protocole SMTP et EHLO pour le protocole ESMTP.

Si nous reprenons l'exemple précédent, voici ce que nous obtenons en réponse du serveur smtp.site2.be lors de l'envoi de la commande EHLO:

```
220 smtp.site2.be SMTP sendmail ready
ehlo smtp.site1.be (ou alors site1.be)
250-smtp.site2.be says hello
250-SIZE 0
250-8BITMIME
250-DSN
250-ETRN
250-AUTH LOGIN CRAM-MD5
250-AUTH=LOGIN
250-EXPN
```

250-SIZE 0 est une extension reprise dans la norme RFC 1427 indiquant qu'il n'y a pas de limite dans la taille du courrier pouvant être accepté par le serveur.

250-8BITMIME est une extension reprise la norme RFC 1426 indiquant que le serveur peut accepter un corps de message codé en ASCII sur 8 bits.

Le serveur indique également qu'il accepte les commandes DSN, ETRN, EXPN

250-AUTH LOGIN CRAM-MD5 est une extension de la norme RFC 2554 qui permet au serveur d'indiquer que le client peut utiliser une méthode d'authentification et de façon optionnelle l'utilisation d'une couche de sécurité sous la forme SASL (Simple Authentication and Security Layer). Dans notre exemple, CRAM-MD5 est une technique d'authentification de type CHALLENGE/REPONSE permettant d'éviter la transmission du mot de passe en clair à travers le réseau. La méthode d'authentification de type LOGIN utilise une technique d'encodage de type Base64 mais qui est triviale et facilement décodable ce qui ne correspond pas à une réelle sécurité dans le transfert du mot de passe. Nous reprendrons un exemple de ces deux techniques ultérieurement.

2. Etude du protocole SMTP.

2.1. Envoi d'un courrier.

Comme repris dans notre exemple précédent, nous allons analyser plus en détail la partie du protocole nous permettant l'envoi d'un simple courrier sans qu'il ne soit question dans cette norme d'une possibilité d'attachement d'un fichier quelconque en utilisant un type MIME.

Une transaction de courrier SMTP se déroule en trois étapes. La transaction est initiée par une commande MAIL donnant l'identification de l'émetteur. Une série contenant une ou plusieurs commandes RCPT suit, donnant les informations sur les destinataires. Puis une commande DATA passe le contenu du message. Dans cette troisième phase, la marque de fin de données à transmettre marque la fin de la transaction.

La première étape de la procédure est donc la commande MAIL. L'argument <route-inverse> contient le nom de la boîte aux lettres de l'émetteur.

```
MAIL <SP> FROM:<reverse-path> <CRLF>
```

Cette commande indique au récepteur SMTP qu'une nouvelle transaction de courrier débute et lui demande de réinitialiser ses tables d'états et ses tampons, y compris ses boîtes de réception et toute donnée de courrier latente. Elle lui donne le chemin inverse qui pourra être utilisé en cas de rapport d'erreur à transmettre. Si l'émetteur accepte la commande, un code 250 OK est renvoyé à l'émetteur.

L'argument <reverse-path> peut contenir plus d'une boîte aux lettres. On peut y inscrire une liste de plusieurs boîtes aux lettres dans des hôtes différents. La première boîte inscrite dans l'argument <reverse-path> doit néanmoins être celle désignant l'émetteur du message.

La deuxième étape de la procédure est l'émission des commandes RCPT.

```
RCPT <SP> TO:<forward-path> <CRLF>
```

Cette commande transmet une adresse de courrier désignant l'adresse de la personne devant recevoir le courrier. Si elle est acceptée, le récepteur SMTP renvoie une réponse de code "250 OK", et mémorise le chemin d'acheminement. Si le récipiendaire est inconnu, le récepteur SMTP renvoie un code d'erreur "550 Failure". Cette seconde étape de la procédure peut être répétée autant de fois que nécessaire.

L'argument <forward-path> peut contenir plus d'une adresse de boîte aux lettres. On peut y inscrire une liste de boîtes aux lettres dans des hôtes différents. Le premier hôte à être mentionné dans l'argument <forward-path> sera l'hôte à qui est envoyé cette commande.

La troisième étape consiste en l'émission de la commande DATA.

DATA <CRLF>

Si elle est acceptée, le récepteur SMTP renvoie une réponse de code "354 Intermediate" et prend en compte toutes les lignes de texte suivantes comme étant le texte du message. Lorsque la marque indiquant la fin du texte est reçue et enregistrée, le récepteur SMTP envoie une réponse de code "250 OK".

2.2. Transfert.

Il existe certains cas où l'information concernant un destinataire donnée dans la < forward-path > est incorrecte, mais le récepteur SMTP connaît la destination exacte. Dans un tel cas, l'une des réponses suivantes pourra être émise pour permettre à l'émetteur de contacter la bonne destination.

```
251 User not local; will forward to <forward-path>
```

Cette réponse indique que le récepteur SMTP sait que la boîte email du destinataire est hébergée sur un autre hôte et indique le chemin d'accès correct de cette boîte aux lettres pour des messages futurs. Notez que l'hôte, ou l'utilisateur ou les deux peuvent être différents de ceux de la requête originale. Le récepteur prend dans ce cas la responsabilité de délivrer le message actuel à la bonne destination.

```
551 User not local; please try <forward-path>
```

Cette réponse indique que le récepteur SMTP sait que la boîte email du destinataire est sur un autre hôte et signale le chemin d'accès correct à utiliser pour joindre cette boîte. Notez que l'hôte, ou l'utilisateur ou les deux peuvent être différents de ceux de la requête originale. Dans ce cas, le récepteur refuse d'accepter des messages pour cet utilisateur, et l'émetteur devra soit rediriger explicitement son courrier conformément à l'information fournie ou arrêter la transaction et retourner un message d'erreur à son utilisateur.

2.3. Vérification d'adresse et expansion de listes.

SMTP propose au titre de fonctionnalités additionnelles, des commandes pour vérifier un nom de destinataire ou pour expander une liste de diffusion. Ces deux opérations peuvent être menées respectivement avec les commandes VRFY et EXPN, qui acceptent toutes deux un argument sous forme de chaîne de caractères. Pour la commande VRFY, la chaîne en argument est un nom d'utilisateur, la réponse à cette commande pouvant inclure le nom complet de cet utilisateur et devant fournir une adresse complètement qualifiée de boîte email pour cet utilisateur. Pour la commande EXPN, la chaîne en argument identifie une liste de diffusion, la réponse multi lignes à cette commande pouvant inclure le nom complet des utilisateurs et DEVANT inclure l'ensemble des boîtes email contenues dans la liste.

La sémantique de l'expression "nom d'utilisateur" est assez floue, cette expression étant employée en parfaite connaissance de cause. Si un hôte implémente la commande VRFY ou EXPN, alors on suppose que l'hôte reconnaît au moins les boîtes email locales en tant que "noms d'utilisateur". Mais un hôte peut utiliser une définition toute autre pour les "noms" de ses utilisateurs.

Sur certains hôtes la distinction entre une liste de diffusion et une série d'alias pour une boîte email unique n'est pas très claire, dans la mesure où les structures de données standard peuvent accueillir ces deux types d'entrées, et qu'il est possible de constituer des listes de diffusion réduites à une seule boîte email. Lorsqu'une requête d'expansion d'une liste de diffusion est lancée, une réponse positive peut être renvoyée si, lorsqu'un message est reçu pour cette liste, ce dernier est transmis simultanément à tous les participants inscrits dans cette liste. Dans tout autre cas, un message d'erreur devra être renvoyé (ex., "550 Ceci est un utilisateur, pas une liste de diffusion"). Lorsqu'une requête de vérification d'un utilisateur est lancée, une réponse positive pourra être donnée si la réponse peut être formée d'une liste ne contenant qu'un seul nom. Dans tout autre cas une erreur devra être renvoyée (ex., "550 Ceci est une liste de diffusion, pas une boîte aux lettres").

Dans le cas d'une réponse multi lignes (réponse courante à une commande EXPN), chaque ligne ne doit spécifier qu'une boîte email et une seule. Dans le cas d'une requête ambiguë, par exemple, "VRFY Dupont", sur un hôte où deux personnes nommées Dupont sont hébergées, la réponse devra être de type "553 utilisateur ambigu".