De technische kant van e-mail op de UvT *

Joost van Baal

27 april 2006

Inhoudsopgave

1 Introductie

1.1 Dit document

Dit document maakt deel uit van de presentatie "De technische aspecten van e-mail" die gegeven werd in Eindhoven op dinsdag 25 april 2006. De presentatie werd georganiseerd door Enosig (http://enosig.org/) voor een algemeen publiek.

De tekst is gebaseerd op een iets langere tekst die hoort bij een presentatie die IT Services organiseerde voor technische ondersteuners op de UvT.

Dank aan Wessel Dankers voor inhoudelijke feedback en LaTeX-tips, aan Mark van Diem voor zijn hulp bij de test-mail aan Lisbeth Driessen voor zinvol commentaar op deze tekst en mijn presentatievaardigheden en aan Ruud van der Velden voor het rapporteren van een vervelende fout in deze tekst.

Het document wordt gepubliceerd op http://non-gnu.uvt.nl/pub/uvt-unix-doc/email-uvt.

1.2 Inhoud en doel

Dit document bestaat uit 2 delen: het eerste deel behandelt verschillende UvT specifieke e-mailzaken, terwijl het tweede deel aan de hand van de route die een mailtje van het internet naar de UvT aflegt, iets van de technische achtergronden zal laten zien.

De situatie op de UvT zal als voorbeeld gebruikt worden voor een algemene flinke e-email-site. Op veel andere plekken zal e-mail op een vergelijkbare manier gebruikt worden. De bedoeling is dat je na het lezen van document een gevoel hebt voor wat mailservers zoal doen, en dat je daardoor beter in staat bent bounce-messages te lezen en andere foutmeldingen te interpreteren.

De tekst gaat ervan uit dat je kunt e-mailen, en dus enig idee hebt over wat e-mail ongeveer is.

2 E-mail op de UvT en e-maildiensten van ITS Unix

2.1 Het centrale mailcluster

Er bestaan 5 centrale Unix mailservers op de UvT, die het mailcluster vormen. Deze machines verwerken ongeveer 50.000 e-mails per dag, voor ongeveer 25.000 lokale e-mailadressen. Ongeveer 100.000 keer per dag wordt er mail aangeboden aan het mailcluster.

2.2 Voorkomen van overlast: virusscanning en spamtagging

Het mailcluster zorgt dat er geen virussen naar UvT-PCs gemaild worden, en werkt eraan om de spam-overlast dragelijk te houden. We zullen iets vertellen over de virusscanning, en iets over de manier waarop we SpamAssassin en greylisting gebruiken om spam-overlast te beperken.

2.2.1 Virusscanning

De UvT mailservers bekijken de inhoud van alle e-mailtjes, en gaan na of die eruit ziet als bekende virussen. Ieder half uur wordt een nieuwe ClamAV database met virus signatures opgehaald. De AMaViS-software voert de scan uit.

Wanneer een virus wordt aangetroffen, wordt het mailtje niet afgeleverd. De verzender krijgt in dit geval *geen* foutmelding terug. In het stuk over sender-spoofing kun je lezen waarom.

2.2.2 Spamtagging

Spam is ongevraagde bulkmail. Op dit moment wordt ongeveer 3% van de verwerkte mail door SpamAssassin als spam gemarkeerd; ongeveer 2000 mailtjes per dag dus.

De UvT mailservers voegen speciale headers toe om verslag te doen van de analyse door SpamAssassin. Wanneer de mail niet als spam herkend wordt, wordt bijvoorbeeld toegevoegd:

```
X-Spam-Cookie: 7336f92f47c42c68766855b5fbae5f910cf8644b
X-Spam-Status: No, hits=-2.6, required=6.3 tests=BAYES_00
X-Spam-Level: -
X-Spam-Flag: No
```

of bijvoorbeeld:

```
X-Spam-Cookie: 7336f92f47c42c68766855b5fbae5f910cf8644b
X-Spam-Status: No, hits=6.1, required=6.3 tests=BAYES_80, HTML_90_100,
        HTML_IMAGE_ONLY_08, HTML_MESSAGE, MIME_HTML_MOSTLY, MPART_ALT_DIFF
X-Spam-Level: ******
X-Spam-Flag: No
```

Wanneer de e-mail wel als spam wordt herkend, worden bijvoorbeeld headers als

```
X-Spam-Cookie: 7336f92f47c42c68766855b5fbae5f910cf8644b
X-Spam-Status: Yes, hits=8.9, required=6.3 tests=BAYES_99, LOCAL_KUB_DOMAIN,
        RCVD_HELO_IP_MISMATCH, RCVD_NUMERIC_HELO
X-Spam-Level: ********
X-Spam-Flag: YES
X-Spam-Report: Content analysis details: (8.9 points, 5.0 required)
        pts rule name description --- ----------------------
        ----------------------------------------------
        2.0 LOCAL_KUB_DOMAIN Mentions obsolete kub.nl-domain
        2.2 RCVD_HELO_IP_MISMATCH Received: HELO and IP do not match, but should
        1.2 RCVD_NUMERIC_HELO Received: contains an IP address used for HELO
        3.5 BAYES_99 BODY: Bayesian spam probability is 99 to 100%
        [score: 1.0000]
```

toegevoegd. Verder wordt in dit geval het subject van de e-mail herschreven: `Subject: Buy OEM Software` wordt vervangen door `Subject: *****SPAM***** Buy OEM Software`.

Dit is zo gedaan zodat met verschillende e-mailclients (waaronder webmail-applicaties) makkelijk gefilterd kan worden op spam. Verder maakt het Spam-Report het mogelijk een vermoeden te krijgen waarom mail als spam gemarkeerd is.

Merk op dat mail die op deze manier als spam herkend wordt, nooit geweigerd wordt door het cluster: centraal worden er alleen headers toegevoegd. De Novell servers en andere machines die de mail uiteindelijk accepteren kunnen dus zelf beslissen wat ze met de spam doen: in een speciale spam-folder zetten bijvoorbeeld, of zo'n beslissing overlaten aan de gebruiker (b.v. met door de gebruiker in te stellen filtering-regels).

2.2.3 Spamtagging en spam-cookie

Verder zie je dat er een Spam-Cookie is toegevoegd aan de headers. Dit wordt gedaan om te voorkomen dat mail onterecht aan (resource-intensieve) spam-analyse wordt onderworpen. (Bijvoorbeeld omdat dat al gedaan is.) Spam-analyse wordt immers alleen uitgevoerd als er nog geen geldig spamcookie in de headers staat.

2.2.4 Greylisting

Greylisting is iets wat plaatsvindt *voordat* de mail daadwerkelijk geaccepteerd wordt: voor iedere nieuwe combinatie van afzender-IP/envelope-from/envelope-to wordt eerst de acceptatie van de mail uitgesteld, en wanneer een aanbiedende mailserver het daarna weer probeert, wordt de mail wel geaccepteerd. Na 5 succesvolle afleverpogingen wordt de verzendende mailserver toegevoegd aan de whitelist database: die server kan dan in het vervolg direct mail afleveren. Dit idee maakt gebruik van een bug in veel spam-software: die gaat vaak niet goed om met uitstel van acceptatie, en legt dat uit als een permanente weigering.

Mail die op deze manier voor de eerste keer wordt aangeboden wordt minimaal 1 minuut vertraagd afgeleverd. Een vertraging van niet-spam mail van meer dan 3 minuten is uitzonderlijk.

Dit is trouwens de reden dat hoewel er 100.000 keer per dag mail wordt aangeboden, er slechts 50.000 e-mails per dag verwerkt worden.

We hebben sterk de indruk dat deze check heel veel spam tegenhoudt.

3 E-mail op het internet: de techniek

3.1 Inleiding

3.1.1 Het centrale mailcluster

Zoals gezegd bestaat het centrale UvT mailcluster uit 5 Unix servers. Mailboxen op de UvT leven niet op deze mailservers, maar (voor het overgrote deel) op de verschillende Novell machines.

Naast de Novell machines en de mailservers zijn voor e-mail nog belangrijk lists.uvt.nl: de Mailman listserver en stuwww.uvt.nl: de machine van SBIT die mail voor de studentenverenigingen ontvangt. ITS Unix beheert 1 machine die daadwerkelijk mailboxen bevat: imap.uvt.nl (die ook webmail.uvt.nl heet).

Hier een plaatje van het centrale Unix mailcluster:

```
                 .---------------------.
novell servers <- |                     |
 imap.uvt.nl   <- |    agnes  herschel  |  mail{1,2}.uvt.nl
                 |                     |  {a,b,c}.mx.uvt.nl
lists.uvt.nl   <- |       germain       |       <---        Internet
stuwww.uvt.nl  <- |                     |        MX
                 `---------------------'
                            ^
                            |
                 .---------------------.
                 |                     |
                 |  piscopia  chatelet |
                 |                     |
                 `---------------------'
                            ^
                            | mailhost.uvt.nl
                            |

            UvT PCs, lists.uvt.nl, stuwww.uvt.nl
```

agnes, herschel, germain, piscopia en chatelet zijn de namen van de 5 mailservers.

De gebruikte software is: Postfix, postgrey, heartbeat, fair (http://non-gnu.uvt.nl/fair), SpamAssassin, amavisd-new en ClamAV.

3.1.2 Een mailtje van het internet naar de UvT

In het volgende plaatje is een voorbeeld weergegeven van de weg die een e-mailtje aflegt, als het vanaf het internet naar een UvT-er gestuurd wordt.

```
    PC
    | SMTP
    v
smtp.xs4all.nl     <----------.
    |                         |
_ _ |     _ _ internet  _ _ _ |      _ _ _ _ _
    | SMTP                    |  DNS
    |                         v
    |                    ns1.uvt.nl
    v
a.mx.uvt.nl <--------> aliases
    |          file
    | SMTP
    v
atommail.uvt.nl
    |
    | IMAP/SSL of (voor webmail) HTTP/SSL
    v
    PC
```

Aan de hand van dit plaatje zal uitgelegd worden hoe internet-mail werkt.

3.2 Het SMTP e-mailprotocol

Wanneer iemand op het internet een mailtje wil sturen naar iemand op de UvT, dan begint dat meestal bij de PC van de verzender: mijn PC thuis, bijvoorbeeld. Ik heb mijn e-mailprogramma verteld dat ik joostvb@xs4all.nl als afzender wil gebruiken. Verder heb ik mijn PC verteld dat alle uitgaande mail naar de mailserver van mijn ISP gestuurd moet worden: smtp.xs4all.nl. Voor dit versturen wordt het SMTP protocol gebruikt.

Het Simple Mail Transfer Protocol (SMTP) is in 1982 (dus al vóórdat het World Wide Web bestond) vastgelegd in RFC 821 (http://www.faqs.org/rfcs/rfc821.html). Overigens: Al in 1971 werd het eerste e-mailtje over een computernetwerk verstuurd (zie http://openmap.bbn.com/~tomlinso/ray/home.html). Later is het protocol herzien, de laatste technische beschrijving (van 2001) staat in RFC 2821 (http://www.faqs.org/rfcs/rfc2821.html).

Deze documenten zijn open Internet standaarden: je hoeft b.v. niemand licentiekosten te betalen als je een mail server zou willen (laten) schrijven. Er zijn dan ook veel verschillende e-mail servers en e-mailprogramma's die met dit protocol om kunnen gaan.

3.3 Een SMTP sessie met telnet

SMTP is een client-server protocol. MS Outlook of Mozilla Thunderbird kunnen als een SMTP client opereren wanneer ze mail versturen.

We kunnen Outlook nadoen, door het telnet-programma te gebruiken. Op deze manier kun je dus een mailtje versturen zonder dat je een echt e-mailprogramma nodig hebt. Dit is bijvoorbeeld handig als je dingen wilt testen. In het voorbeeld wordt telnet vanaf een Linux-systeem gebruikt. De telnet-client die met PuTTY voor MS Windows wordt meegeleverd, is echter ook prima bruikbaar.

Een SMTP telnet-sessie ziet er als volgt uit. Voor de duidelijkheid heb ik de regels die de server teruggeeft laten inspringen, zodat die goed te onderscheiden zijn van de regels die ik typ.

```
joostvb@gelfand:~% telnet smtp.xs4all.nl 25
Trying 194.109.6.51...
Connected to smtp.xs4all.nl.
Escape character is '^]'.
  220 smtp-vbr1.xs4all.nl ESMTP Sendmail 8.13.6/8.13.6; Mon, 24 Apr 2006
     10:17:27 +0200 (CEST)
EHLO gelfand.mdcc.cx
  250-smtp-vbr1.xs4all.nl Hello gelfand.mdcc.cx [80.126.189.155], pleased
     to meet you
  250-ENHANCEDSTATUSCODES
  250-PIPELINING
  250-8BITMIME
  250-SIZE 31457280
  250-DSN
  250-AUTH LOGIN PLAIN
  250-DELIVERBY
  250 HELP
MAIL FROM:<joostvb@xs4all.nl>
  250 2.1.0 <joostvb@xs4all.nl>... Sender ok
RCPT TO:<M.H.P.vDiem@uvt.nl>
  250 2.1.5 <M.H.P.vDiem@uvt.nl>... Recipient ok
DATA
  354 Enter mail, end with "." on a line by itself
From: Joost van Baal <joostvb@xs4all.nl>
To: Mark van Diem <M.H.P.vDiem@uvt.nl>
Subject: test, negeer aub
Message-Id: <20060424.123@gelfand.mdcc.cx>
Date: Mon, 24 Apr 2006 10:18:52 +0200

Hoi Mark,

Dit is een testje voor mijn e-mail-praatje van vrijdag.

Groeten, Joost
.
  250 2.0.0 smtp-vbr1.xs4all.nl accepted message k308HRt7073664
QUIT
  221 2.0.0 smtp-vbr1.xs4all.nl closing connection
Connection closed by foreign host.
```

3.4 Sender-spoofing; envelope, header en body

Laten we eens in detail naar bovenstaande sessie kijken.

Ik maak een verbinding met TCP poort 25. Dat is de standaard TCP netwerkpoort waarop e-mail servers luisteren.

Ik maak mezelf met het "EHLO" commando bekend als SMTP client gelfand.mdcc.cx. Sommige servers checken of clients zich met een juiste naam bekend maken, door te kijken of het IP adres dat de TCP sessie opzet, overeenkomt met de EHLO naam.

De server geeft regels terug die beginnen met "250". Hierin vertelt de server welke SMTP extensies ondersteund worden. De server vertelt ook dat hij e-mails van maximaal 30 MB accepteert.

Daarna vertel ik dat ik mail wil versturen met joostvb@xs4all.nl als afzenderadres. Merk op dat de UvT-server op dit punt alleen mijn IP-adres weet. De server heeft geen enkele manier om na te gaan of het echt wel Joost van Baal, de eigenaar van de mailbox achter joostvb@xs4all.nl, is die de client op deze IP controleert. Jij kunt dus ook best mail met mijn afzender-adres sturen, of mail met afzenderadress president@whitehouse.gov. Verder kan het mailtje met afzenderadres president@whitehouse.gov dat jij krijgt, dus best door Joost van Baal verstuurd zijn. Je hoeft trouwens helemaal geen telnet te gebruiken om mail met vervalste afzender te versturen: je kunt gewoon de instellingen van je Outlook aanpassen, en daar president@whitehouse.gov als afzenderadres instellen. Dit is trouwens de reden dat het mailcluster geen foutmelding terug kan geven als er een virus in een mailtje is aangetroffen: virus-mailtjes gebruiken vrijwel altijd mailadressen van onschuldige derden als vervalst afzender-adres.

Ik vertel dat de mail naar M.H.P.vDiem@uvt.nl gestuurd moet worden.

Dan kondig ik met “DATA” aan dat ik de werkelijke e-mail wil aanbieden. De server zegt dat ik de tekst moet beëindigen met een regel die alleen uit een punt bestaat. De gegevens die ik hiervoor heb gegeven (EHLO, MAIL FROM, RCPT TO) heten de envelope. Alles wat na DATA komt heet de mailheader en mailbody. Header en body worden gescheiden door een lege regel.

Ik typ nu de e-mail, inclusief de headers. Als ik hier had getypt: `To: The President <president@whitehouse.gov>`, dan was die mail toch naar M.H.P.vDiem@uvt.nl gestuurd: de mailserver kijkt immers alleen naar de envelope (RCPT TO:).

Ik sluit af met een punt. De mailserver vertelt me dan dat ie het mailtje geaccepteerd heeft, en dat ie het van mailqueue-ID `k308HRt7073664` heeft voorzien. Zo’n queue-ID is handig als de beheerder van de ontvangende mailserver later uit wil vinden wat er met het mailtje gebeurd is.

Ik sluit de sessie af met QUIT, en ook de server hangt op.

Merk op dat ik dus helemaal geen echt werkend e-mailadres, of een e-mailprogramma of een mailbox nodig heb om mail te versturen. Voor het versturen van e-mail heb je bijna niks nodig. (Voor het *ontvangen* van e-mail heb je dat allemaal *wel* nodig.)

3.5 SMTP naar het mailcluster; DNS, `Received:`-headers

De mail is nu aanbeland op de machine smtp.xs4all.nl (die zichzelf ook wel smtp-vbr1.xs4all.nl noemt.) Die machine zal vervolgens de mail naar een UvT-server doorsturen. Om erachter te komen welke machine gebruikt moet worden voor mail voor @uvt.nl, wordt DNS gebruikt. De xs4all-machine vraagt de DNS servers naar het zogenaamde MX-record voor het uvt.nl maildomein. Die vraag kunnen we zelf ook stellen:

```
joostvb@gelfand:~% dig uvt.nl mx
[...]
;; ANSWER SECTION:
uvt.nl.                 85930   IN      MX      10 c.mx.uvt.nl.
uvt.nl.                 85930   IN      MX      10 a.mx.uvt.nl.
uvt.nl.                 85930   IN      MX      10 b.mx.uvt.nl.
[...]
```

Het antwoord zal gegeven worden door bijvoorbeeld ns1.uvt.nl. (Onder MS Windows kun je dit met het programma nslookup doen.)

Tegelijk met RFC 821 werd RFC 822 gepubliceerd, waarin beschreven werd hoe de inhoud van een e-mailbericht eruit moet zien: wat moet er bijvoorbeeld in de headers staan. Een laatste versie van deze open standaard is gepubliceerd in RFC 2822 (http://www.faqs.org/rfcs/rfc2822.html). Ook de rol van headers die met een `X-` beginnen staat daarin beschreven. (`Received:`-headers staat beschreven in RFC 2821).

Wanneer bekend is welke server gebruikt kan worden, start de xs4all-machine de SMTP-sessie; de antwoorden die de mailcluster-machine geeft zijn ingesprongen. Merk trouwens op dat de rol van de xs4all-machine nu is omgedraaid: in de vorige sessie speelde die server, nu speelt ie SMTP-client.

```
[connectie vanaf smtp-vbr1.xs4all.nl met a.mx.uvt.nl op poort 25]
 220 agnes.uvt.nl ESMTP Debian GNU/Linux Postfix
EHLO smtp-vbr1.xs4all.nl
 250-agnes.uvt.nl
 250-PIPELINING
 250-SIZE 10240000
 250-VRFY
 250-ETRN
 250 8BITMIME
MAIL FROM:<joostvb@xs4all.nl>
 250 Ok
RCPT TO:<M.H.P.vDiem@uvt.nl>
 250 Ok
DATA
 354 End data with <CR><LF>.<CR><LF>
Received: from gelfand.mdcc.cx (gelfand.mdcc.cx [80.126.189.155])
        by smtp-vbr1.xs4all.nl (8.13.6/8.13.6) with ESMTP id k308HRt7073664
        for <M.H.P.vDiem@uvt.nl>; Mon, 24 Apr 2006 10:17:51 +0200 (CEST)
        (envelope-from joostvb@xs4all.nl)
From: Joost van Baal <joostvb@xs4all.nl>
To: Mark van Diem <M.H.P.vDiem@uvt.nl>
Subject: test, negeer aub
Message-Id: <20060424.123@gelfand.mdcc.cx>
Date: Mon, 24 Apr 2006 10:18:52 +0200
X-Virus-Scanned: by XS4ALL Virus Scanner

Hoi Mark,

Dit is een testje voor mijn e-mail-praatje van vrijdag.

Groeten, Joost
.
 250 Ok: queued as D82931623
QUIT
 221 Bye
```

Merk op dat de xs4all-machine een extra `Received:`-header aan de mail heeft geplakt; in die header heeft deze machine ook verteld wat de envelope in de vorige sessie was. Verder heeft deze machine blijkbaar gescand voor virussen: dat geeft ie weer in een extra header.

3.6 Soorten e-mailadressen in headers en envelopes, bounce-messages

We zagen al dat er typisch 4 mailadressen betrokken zijn bij 1 SMTP-sessie: de envelope-from (`MAIL FROM:`), de envelope-to (`RCPT TO:`), de header-from (`From: `), en de header-to (`To: `).

De envelope-from wordt door de ontvangende server gebruikt om foutmeldingen heen te sturen, wanneer er na het beëindigen van de sessie nog iets fout gaat. Het zou bijvoorbeeld kunnen dat de mailbox van de geadresseerde overvol is, of dat het mailadres niet bestaat: User unknown.

Zo'n fout komt typisch tijdens een volgende SMTP-sessie aan het licht: een server die een alias-file heeft geeft in de sessie een foutcode die met dit probleem overeenkomt. De server van de ISP verpakt dat bericht van de mailrelay van de UvT dan in een bounce-message die naar de envelope-from gestuurd wordt.

Hieronder zie je het genereren van een bounce-message in een schemaatje. De gebruiker stuurt vanaf zijn PC een mailtje voor bestaatniet@uvt.nl naar de mailrelay van de ISP. Daarna vertelt de mailrelay van de ISP tegen de mailrelay van de UvT dat hij een bericht voor bestaatniet@uvt.nl heeft:

```
PC     ->                 mailrelay ISP    ----------->    mailrelay UvT

                                 MAIL FROM:<joostvb@xs4all.nl>
                                 RCPT TO:<bestaatniet@uvt.nl>
```

De UvT mailrelay reageert met een foutmelding:

```
                          mailrelay ISP    <-----------    mailrelay UvT

                            550 <bestaatniet@uvt.nl>: Recipient address
                          rejected: User unknown in local recipient table
```

De ISP mailrelay maakt hiervan een bounce-message, die hij aan de gebruiker aanbiedt (b.v. door het naar de POP-server te sturen, waar de gebruiker de mail later vanaf kan halen.)

```
 POP-server <----------- mailrelay ISP

MAIL FROM:<>
RCPT TO:<joostvb@xs4all.nl>

DATA

From: Mail Delivery System <MAILER-DAEMON@xs4all.nl>
Subject: Undelivered Mail Returned to Sender
To: joostvb@xs4all.nl

I'm sorry to have to inform you that your message could
not be be delivered to one or more recipients.

 <bestaatniet@uvt.nl>: host c.mx.uvt.nl[137.56.247.20]
   said: 550 <bestaatniet@uvt.nl>: Recipient address
   rejected: User unknown in local recipient table (in
   reply to RCPT TO command)
```

De envelope-to wordt gebruikt om de mailbox van de geaddresseerde te vinden.

De header-to en de header-from zijn er *alleen maar* voor de uiteindelijke ontvanger. Over het algemeen kijken mailservers daar helemaal niet naar. Je e-mailprogramma kijkt er trouwens wel naar: als je op "Reply" klikt, dan wordt er een mailtje gestuurd naar het adres dat in de Reply-To of From-header van het origele mailtje stond.

Wanneer je iemand Bcc-t op een mailtje, dan zal het adres van die persoon in de envelope-to staan, en niet in de e-mail headers: de software verwijdert de Bcc-header voordat ie het mailtje aflevert.

3.7 Verkeerde envelope-from bij e-mails die door programma's worden verstuurd

Een veel voorkomend probleem met envelope-from adressen komt aan het licht bij CGIwebapplicaties. Typisch vragen webapplicaties om het mailadres van de gebruiker. Daarna sturen ze e-mail naar dat adres. Wanneer het gebruikte mailadres onjuist is (bijvoorbeeld omdat de gebruiker een typfout maakte), dan zal het mailcluster die mail niet kunnen afleveren. Er zal dus een bounce-message gegenereerd worden, waarin dit probleem wordt weergegeven. Nu is het zo dat veel

Unix-servers voor e-mail die door de lokale webserver wordt gegenereerd (b.v. via een CGI-script) het adres www-data@hostnaam.uvt.nl als envelope-from gebruiken. Die bounce-message zal het cluster dus naar www-data@hostnaam.uvt.nl proberen te sturen. Ook *dat* kan foutgaan. Immers, mogelijk is TCP poort 25 van hostnaam.uvt.nl niet bereikbaar vanaf het mailcluster. Wanneer die poort *wel* bereikbaar is, kan het dat niemand de mail voor de gebruiker www-data op die machine leest.

Het is dus beter om een functioneel adres dat gebruikt wordt door de beheerder van de webapplicatie, als envelope-from te gebruiken. Het hangt van de specifieke webapplicatie af hoe je dit voor elkaar kunt krijgen. (Op Unix-systemen gaat het over het algemeen door de `-f`-vlag van `sendmail(1)` te gebruiken.) Let vooral op dat je bounces *niet* bij onschuldige derden (zoals www-data@uvt.nl) terecht komen!

3.8 Vervalste `Received:`-headers

Mailheaders worden vaak vervalst door spammers: bijvoorbeeld

```
Received: from 24.32.64.156 (unknown [218.146.9.28])
        by agnes.uvt.nl (Postfix) with SMTP id A2C50401
        for <postmaster@kub.nl>; Tue, 18 Apr 2006 09:44:35 +0200 (CEST)
Received: from mail2.tda-inc.com (mail2.tda-inc.com [65.219.50.26]) by
        smtp.montreal.com with esmtp; abr, 18 2006 3:39:24 -0100
```

en

```
Received: from 163.66.78.222.broad.dynamic.ly.fj.cn.cndata.com (unknown
        [222.78.69.163]) by agnes.uvt.nl (Postfix) with SMTP id 2996143D
        for <postmaster@tias.be>; Tue, 18 Apr 2006 09:29:22 +0200 (CEST)
X-Message-Info: QNMnVU21fBtSChbh6s541M285NWTbbkUCZjt
Received: from t15.caramail.com (75.176.237.108) by mfs508-dim.i-frane.com
        with Microsoft SMTPSVC(5.0.2195.6824);
        Tue, 18 Apr 2006 07:28:21 -0100
Received: from facecatapulteatenf3 (trellis168.242.98.211)
        by loja.net (fodrew5) with SMTP id <743985515997940dt9qr>
        (Authid: OpheliaHarrington); Tue, 18 Apr 2006 14:21:21 +0600
```

Om goed `Received:`-headers te kunnen lezen, moet je weten hoe je eigen mailinfrastructuur in elkaar zit, en welke headers je dus redelijkerwijs kunt vertrouwen. Over het algemeen kun je alleen de headers die door je eigen servers zijn toegevoegd vertrouwen.

3.9 SMTP naar de Novell-machine; afleveren van de e-mail in de mailbox

Het mailtje van Joost aan Mark staat inmiddels in de mailqueue van het mailcluster. De mailclustermachine kijkt in het UvT alias-file, en ziet daar staan: `M.H.P.vDiem:  mvd@atommail.uvt.nl`. De mailcluster-machine kan het mailtje nu dus aanbieden aan Novell-server atommail.uvt.nl, die het mailtje kan afleveren in de mailbox van de ontvanger:

```
[connectie vanaf agnes.uvt.nl met atommail.uvt.nl op poort 25]
 220 atommail.uvt.nl Novonyx SMTP ready $Revision: 10195 $
EHLO agnes.uvt.nl
 250-atommail.uvt.nl Pleased to meet you
 250-ETRN
 250-STARTTLS
 250-AUTH LOGIN
 250-AUTH=LOGIN
 250-HELP
```

```
 250-EXPN
 250-PIPELINING
 250-8BITMIME
 250-DSN
 250 SIZE 10485760
MAIL FROM:<joostvb@xs4all.nl>
 250 Sender OK
RCPT TO:<mvd@atommail.uvt.nl>
 250 Recipient OK
DATA
 354 Send message, end with <CRLF>.<CRLF>
Received: from localhost (localhost [127.0.0.1])
        by agnes.uvt.nl (Postfix) with ESMTP id D04331653
        for <mvd@atommail.uvt.nl>; Mon, 24 Apr 2006 10:19:15 +0200 (CEST)
Received: from agnes.uvt.nl ([127.0.0.1])
        by localhost (agnes [127.0.0.1]) (amavisd-new, port 10024) with ESMTP
        id 29904-02 for <mvd@atommail.uvt.nl>;
        Mon, 24 Apr 2006 10:19:15 +0200 (CEST)
Received: from smtp-vbr1.xs4all.nl (smtp-vbr1.xs4all.nl [194.109.24.21])
        by agnes.uvt.nl (Postfix) with ESMTP id AB540EEA
        for <M.H.P.vDiem@uvt.nl>; Mon, 24 Apr 2006 10:19:15 +0200 (CEST)
Received: from gelfand.mdcc.cx (gelfand.mdcc.cx [80.126.189.155])
        by smtp-vbr1.xs4all.nl (8.13.6/8.13.6) with ESMTP id k3O8HRt7073664
        for <M.H.P.vDiem@uvt.nl>; Mon, 24 Apr 2006 10:17:51 +0200 (CEST)
        (envelope-from joostvb@xs4all.nl)
From: Joost van Baal <joostvb@xs4all.nl>
To: Mark van Diem <M.H.P.vDiem@uvt.nl>
Subject: test, negeer aub
Message-Id: <20060424.123@gelfand.mdcc.cx>
Date: Mon, 24 Apr 2006 10:18:52 +0200
X-Virus-Scanned: by XS4ALL Virus Scanner
X-Virus-Scanned: by amavisd-new-20030616-p10 (Debian) at uvt.nl
X-Spam-Cookie: 120d18aca32a7adc27397f1cec6d1f77ac8729d6
X-Spam-Status: No, hits=-1.5, required=6.3 tests=BAYES_00
X-Spam-Level: -
X-Spam-Flag: No

Hoi Mark,

Dit is een testje voor mijn e-mail-praatje van vrijdag.

Groeten, Joost
.
 250 OK
QUIT
 221 atommail.uvt.nl So long, and thanks for all the fish
```

Merk op dat het mailcluster nu mvd@atommail.uvt.nl als envelope-to gebruikt (maar niet als header-to): deze envelope-to wijzigt dus tijdens het pad door de verschillende mailservers. Verder heeft het cluster er 3 extra `Received:`-headers aan vastgeplakt. Als laatste is er 1 **`X-Virus-Scanned:`**-header toegevoegd, en 4 **`X-Spam`**-headers.

De Novell server plakt er nu nog als bovenste headers

```
Return-Path: joostvb@xs4all.nl
Received: from agnes.uvt.nl not authenticated [137.56.247.33]
        by atommail.uvt.nl with NetMail SMTP Agent $Revision: 10195 $
        on Novell NetWare;
        Mon, 24 Apr 2006 10:19:16 +0200
```

aan vast, en levert het mailtje af in de mailbox van de ontvanger. Die kan het dan met een e-mailclient met POP of IMAP over SSL of via een webbrowser met HTTP over SSL uit de mailbox halen en lezen.

Merk op dat de verschillende envelope-from's en envelope-to's niet zichtbaar zijn in het mailtje dat uiteindelijk in de mailbox van de gebruiker terecht komt. Sommige servers geven in hun `Received:`-headers (een gedeelte van) de envelope-informatie weer. Maar gegarandeerd is dat allemaal niet. Wat *wel* gegarandeerd is, is de allerlaatst (dus helemaal bovenaan) toegevoegde `Return-Path:`-header; hierin staat de laatste envelope-from.

Wanneer Mark nu op de Reply-knop van zijn e-mailprogramma drukt, een antwoord schrijft, en op de Send-knop duwt, dan zal zijn e-mailprogramma een SMTP-sessie gaan starten met mailhost.uvt.nl: dat is immers de machine-naam die UvT PCs gebruiken om hun uitgaande mail heen te sturen.

3.10 E-mail logfiles

Alle servers waar het mailtje overheen is gegaan leggen in hun logfiles iets vast. Over het algemeen is dat: tijdstip, naam en IP adres van verzendende machine, envelope-from adres tijdens inkomende SMTP-sessie, grootte van het mailtje, Message-ID van het mailtje, eigen queueid, naam en IP adres van ontvangende machine, envelope-to adres tijdens uitgaande SMTP-sessie en het bericht dat ontvangende machine gaf bij het accepteren van het mailtje (b.v. queueid).

Het `Subject:` van de mail wordt over het algemeen niet vastgelegd. Ook wordt er over het algemeen geen copie van de body bewaard.

Hieronder staat een stukje van wat 1 van de betrokken mailcluster-systemen vastlegde.

```
Apr 24 10:19:15 agnes postfix/smtpd[29902]: AB540EEA:
 client=smtp-vbr1.xs4all.nl[194.109.24.21]
Apr 24 10:19:15 agnes postfix/cleanup[29775]: AB540EEA:
 message-id=<20060424.123@gelfand.mdcc.cx>
Apr 24 10:19:15 agnes postfix/qmgr[20008]: AB540EEA: from=<joostvb@xs4all.nl>,
 size=759, nrcpt=1 (queue active)
Apr 24 10:19:15 agnes postfix/smtp[28232]: D04331653: to=<mvd@atommail.uvt.nl>,
 relay=atommail.uvt.nl[137.56.12.145], delay=0, status=sent (250 OK)
```

4 Meer informatie en contact

Een manier om de gevaren van sender-spoofing te lijf te gaan is het gebruik van PGP (zie o.a. http://mdcc.cx/gnupg/).

Naast de anti-spam maatregelen die op de UvT genomen worden is er ook nog SPF: Op dit moment wordt door de internet-gemeenschap over deze techniek om o.a. spam-overlast te bestrijden nagedacht (zie http://en.wikipedia.org/wiki/Sender_Policy_Framework).

MIME (http://en.wikipedia.org/wiki/MIME) is de naam van de techniek die het mogelijk maakt om bijlagen in je e-mails te versturen.

De auteur is te bereiken via joostvb+email-uvt-article@uvt.nl.

`$Id: email-uvt.tex 10195 2006-04-27 14:24:19Z joostvb $`