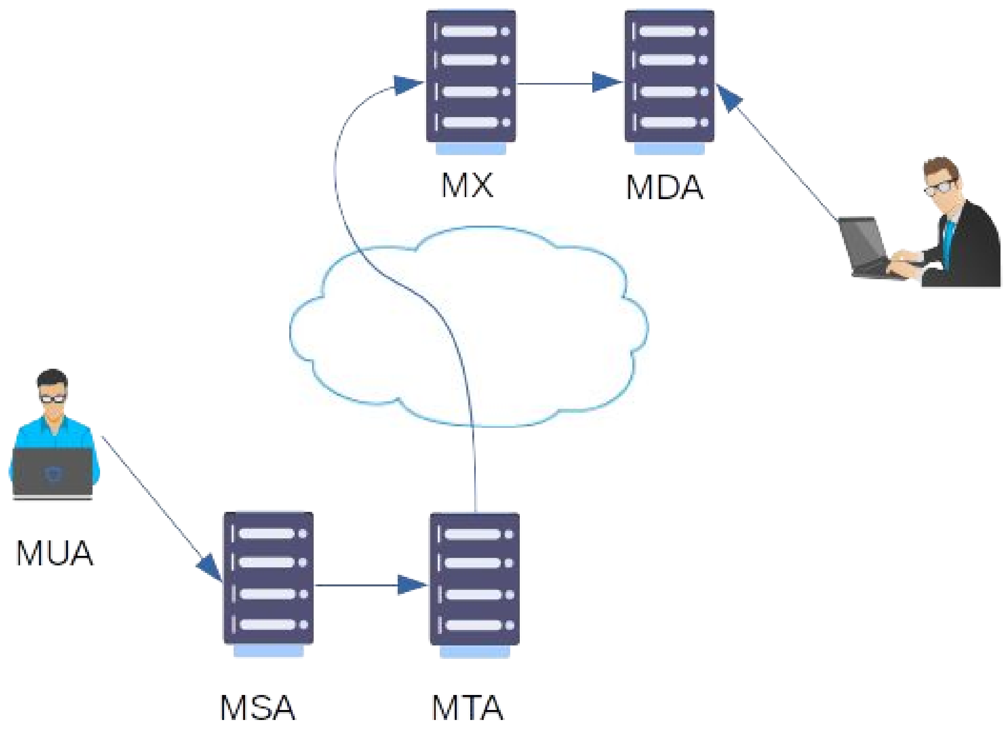
SMTP: how your emails are transported

The email journey

SMTP (**S**imple **M**ail **T**ransport **P**rotocol) is an Internet standard communication protocol for email transmission. SMTP describes how your emails are sent from your email client on your PC or phone, or your webmail interface, to their destination.

The complete journey of an email can be represented as follows:



Some definitions:

- **MUA:** Mail User Agent: a mail client like Outlook, Thunderbird or Gmail, that the user uses to write and send his emails;
- **MSA**: Mail Submission Agent: the mail server part which communicates with the MTA, usually on TCP port 25 or 587 when using encryption, and delivers the mail to the MTA;
- **MTA**: Mail Transfer Agent: the mail server part which will determine where to send the email to the addressee. The MTA will lookup the DNS MX record for the domain part of the address of the recipient to find where to send the email;
- **MX**: Mail eXchange: the destination MTA;
- **MDA**: Mail Delivery Agent: the component that will store the received email in the recipient's mailbox.

Note that several components can actually share the same server.

All the transmissions between the different agents happen via the SMTP protocol. The actual retrieval of the email by the recipient, though, is made by other means we will explain in future articles.

The email transmission language

Definition

As its name implies, SMTP is a *protocol*, i.e. a set of standardized commands and responses between the components so they can understand each others.

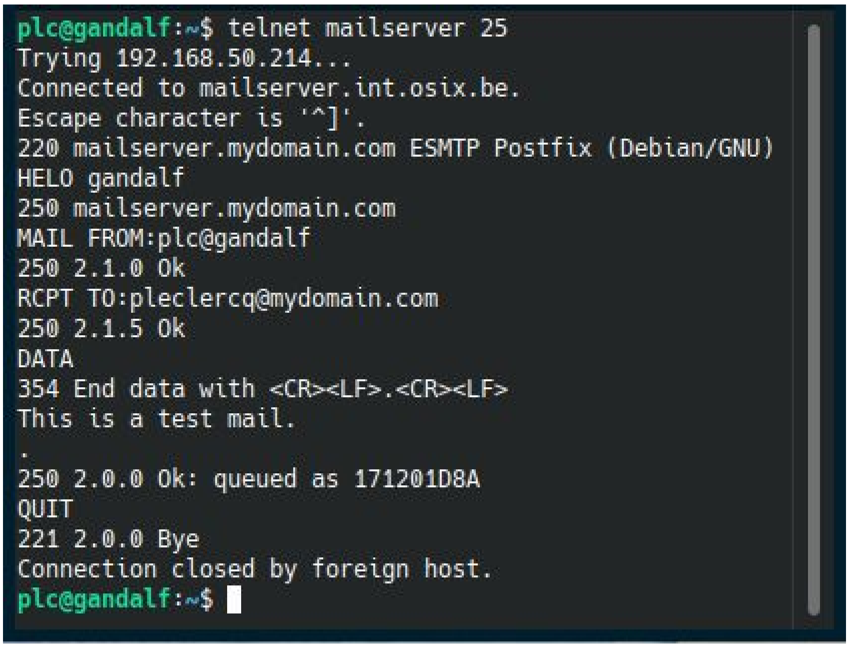
Seen from the MUA (the mail agent), this protocol is really very simple (as once again its name implies).

The protocol is based on 5 main commands:

- **`HELO`**: used by the client to identify itself;
- **`MAIL FROM`**: used by the client to identify the sender;
- **`RCPT TO`**: used by the client to identify the recipient;
- **`DATA`**: introduces the real content of the email. The end of the data is marked by a line only containing a sigle dot (".").
- **`QUIT`**: closes the conversation.

Example

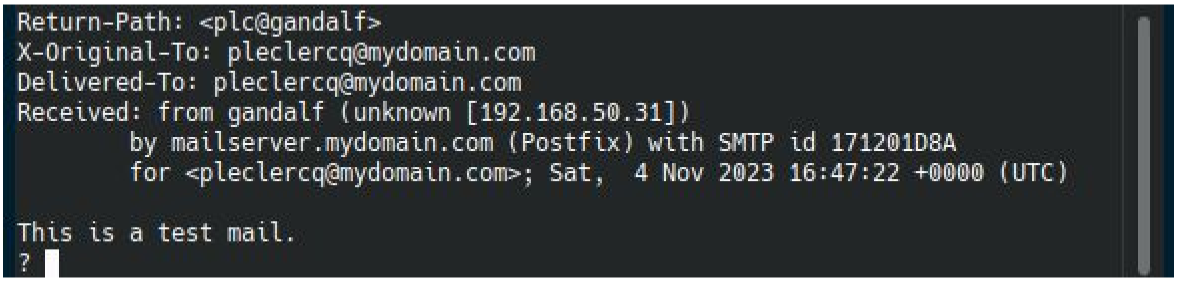
In the following example, we have simplified the infrastructure to the maximum. We have one client machine (*gandalf*) sending a mail to a user having an account on the mail server of the *mydomain.com* domain itself (*mailserver*). We will not use a real mail client; we will directly connect to the SMTP server port of the server and will type the SMTP commands manually.



```
plc@gandalf:~$ telnet mailserver 25
Trying 192.168.50.214...
Connected to mailserver.int.osix.be.
Escape character is '^]'.
220 mailserver.mydomain.com ESMTP Postfix (Debian/GNU)
HELO gandalf
250 mailserver.mydomain.com
MAIL FROM:plc@gandalf
250 2.1.0 Ok
RCPT TO:pleclercq@mydomain.com
250 2.1.5 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
This is a test mail.
.
250 2.0.0 Ok: queued as 171201D8A
QUIT
221 2.0.0 Bye
Connection closed by foreign host.
plc@gandalf:~$
```

The `HELO`, `MAIL FROM`, `RCPT TO`, `DATA`, `QUIT` commands and the lines between "`354...`" and "`250 2.0.0 OK...`" are typed from the client. These are exactly what your usual mail client transmits when you send a mail. The answers from the mail server begin with 3 digits.

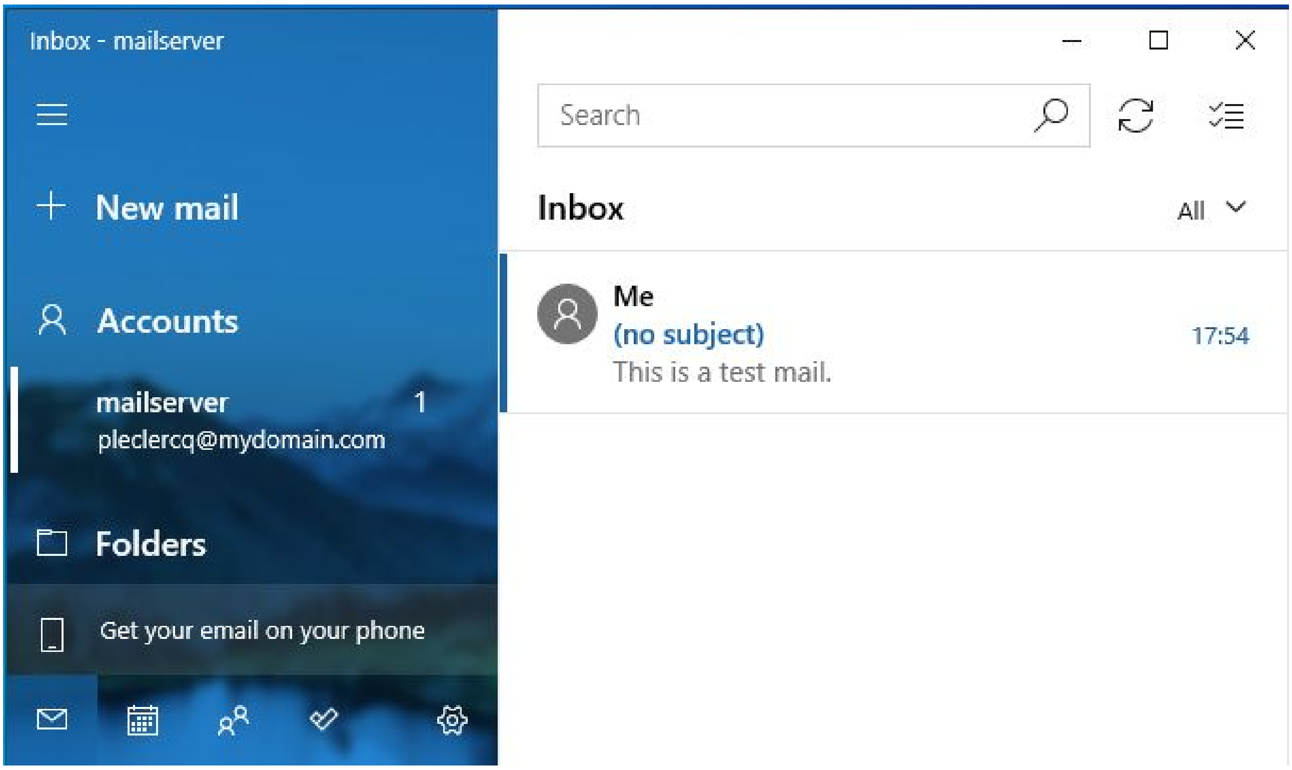
Using a simple character based mail client on the *mailserver* server, we can read the following email content:



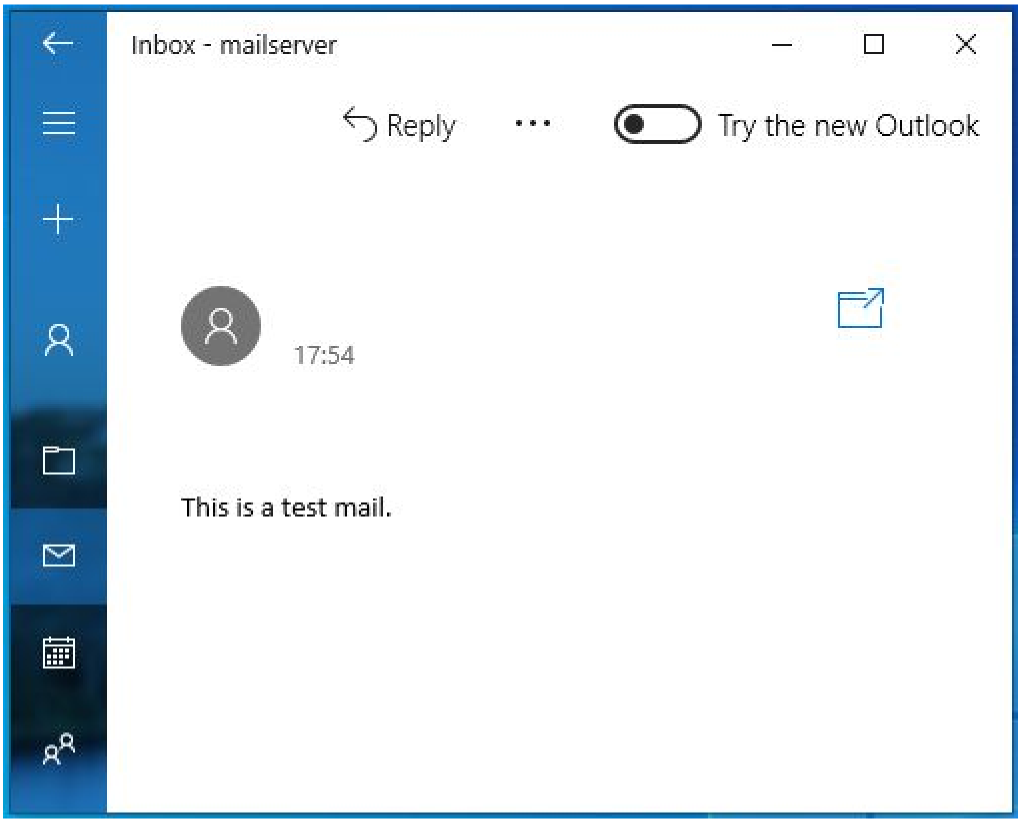
```
Return-Path: <plc@gandalf>
X-Original-To: pleclercq@mydomain.com
Delivered-To: pleclercq@mydomain.com
Received: from gandalf (unknown [192.168.50.31])
        by mailserver.mydomain.com (Postfix) with SMTP id 171201D8A
        for <pleclercq@mydomain.com>; Sat,  4 Nov 2023 16:47:22 +0000 (UTC)

This is a test mail.
?
```

The data above the actual content text of the mail is called the ***mail headers***. They give important technical details about the source and transmission path of the email. We will analyze these in a future article.

Using a graphical mail client (the `mail` application on Windows), this is what the received mail looks like:



Me
(no subject) 17:54
This is a test mail.



Inbox - mailserver

17:54

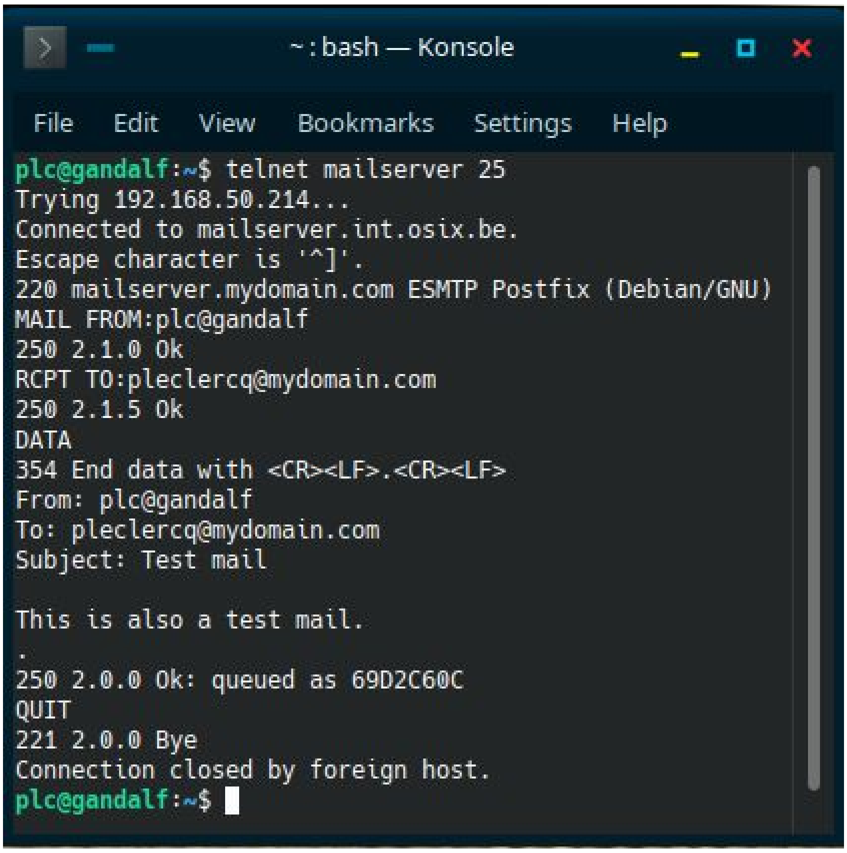
This is a test mail.

Adding more details

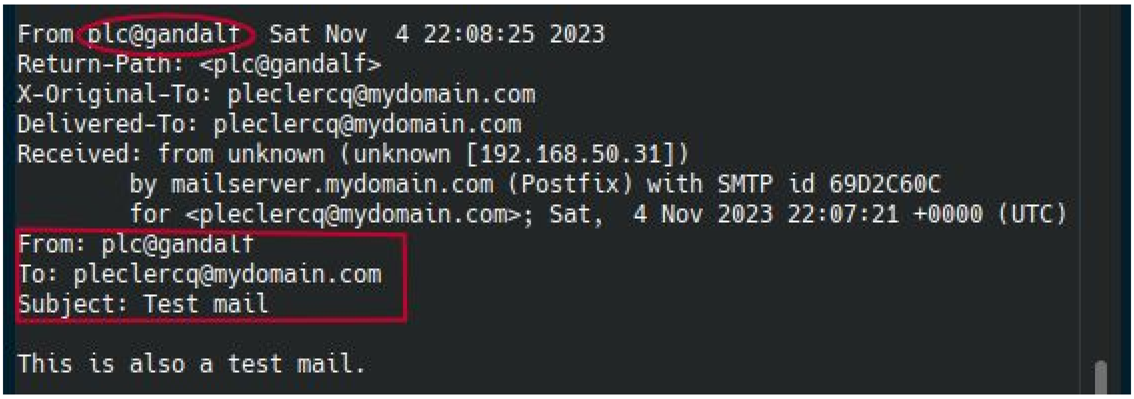
As you can see above, the mail clients do not show a lot of details with only the minimum fields of the email filled.

Several Internet standards define the content of an email. Notably, the fields `From:`, `To:` and `Subject:` are defined and are used by mail clients to enrich the description of the email.

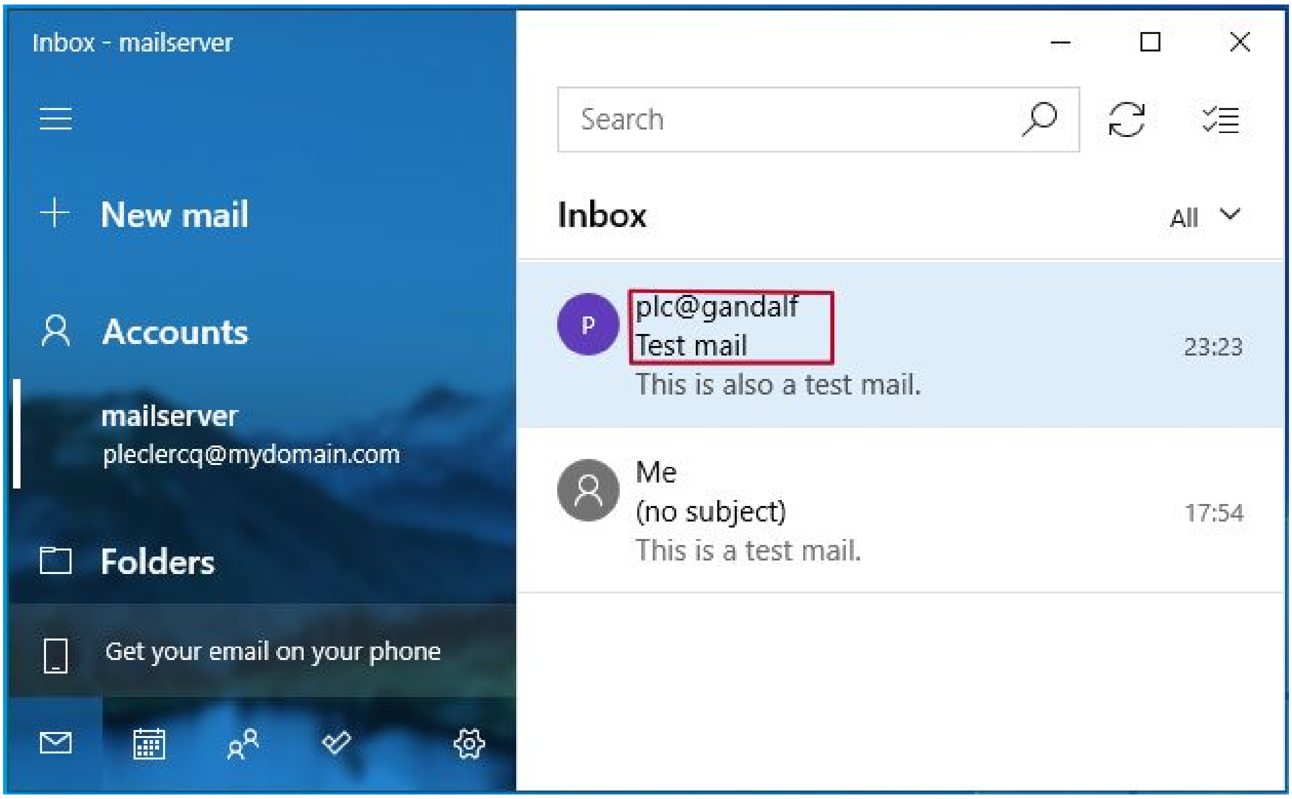
Here is an example of a more complete email and how it appears.



```
plc@gandalf:~$ telnet mailserver 25
Trying 192.168.50.214...
Connected to mailserver.int.osix.be.
Escape character is '^]'.
220 mailserver.mydomain.com ESMTP Postfix (Debian/GNU)
MAIL FROM:plc@gandalf
250 2.1.0 Ok
RCPT TO:pleclercq@mydomain.com
250 2.1.5 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
From: plc@gandalf
To: pleclercq@mydomain.com
Subject: Test mail

This is also a test mail.
.
250 2.0.0 Ok: queued as 69D2C60C
QUIT
221 2.0.0 Bye
Connection closed by foreign host.
plc@gandalf:~$
```



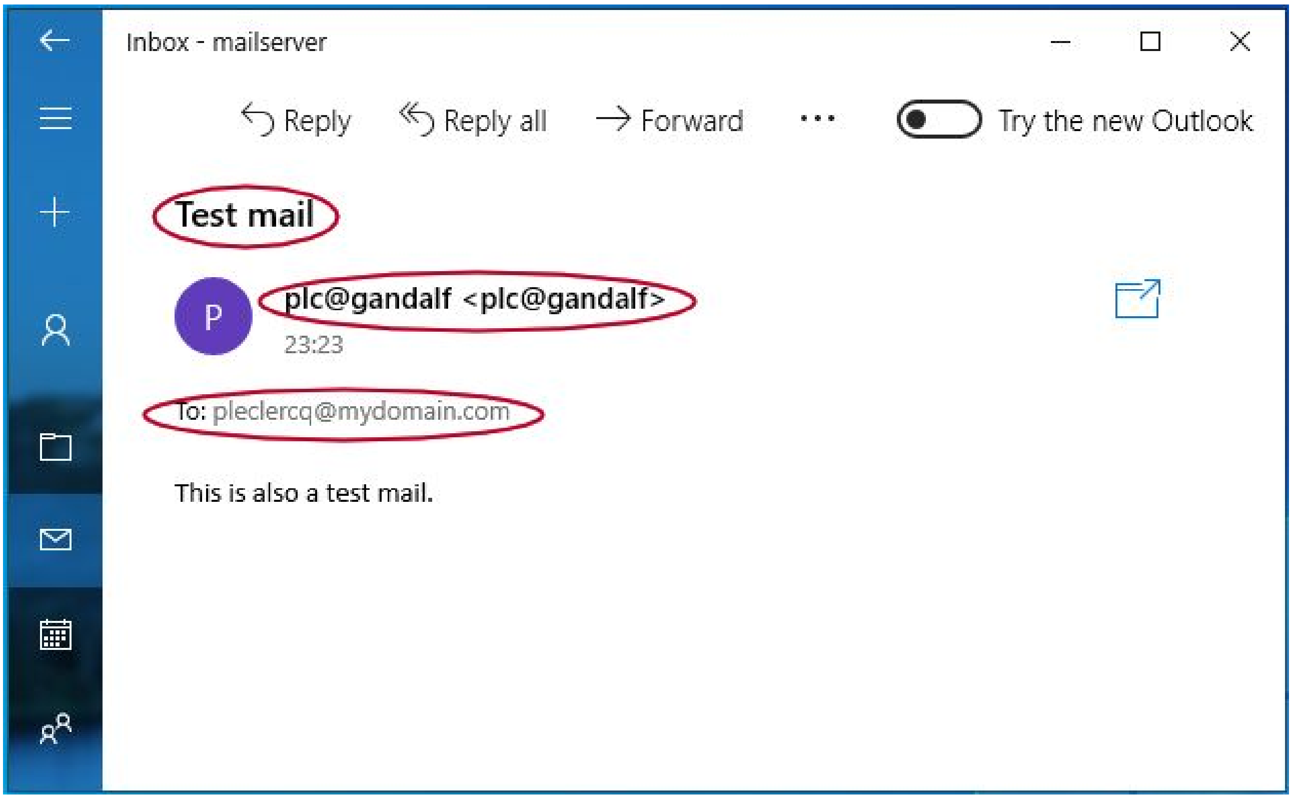
```
From plc@gandalf  Sat Nov  4 22:08:25 2023
Return-Path: <plc@gandalf>
X-Original-To: pleclercq@mydomain.com
Delivered-To: pleclercq@mydomain.com
Received: from unknown (unknown [192.168.50.31])
        by mailserver.mydomain.com (Postfix) with SMTP id 69D2C60C
        for <pleclercq@mydomain.com>; Sat,  4 Nov 2023 22:07:21 +0000 (UTC)
From: plc@gandalf
To: pleclercq@mydomain.com
Subject: Test mail

This is also a test mail.
```



mailserver
pleclercq@mydomain.com

plc@gandalf
Test mail 23:23
This is also a test mail.

Me
(no subject) 17:54
This is a test mail.



Test mail

plc@gandalf <plc@gandalf>
23:23

To: pleclercq@mydomain.com

This is also a test mail.

You can now clearly see the sender name (`plc@gandalf`), the recipient name (`pleclercq@mydomain.com`) and the subject.

Most mail clients actually add these fields when transmitting an email. Usually, when you configure them, they ask how you want your name to be displayed ("display name") to fill the `From:` data field and they copy the `RCPT TO` header field into the `To:` data field.

In a future article, we will see how the existence of these data fields and the way they are displayed in a mail agent can sometimes be misleading and be used by attackers to confuse the recipient.