National Cyber Security Centre
Ministry of Security and Justice

Secure the connections of mail servers

STARTTLS and DANE protect email traffic on the internet

Factsheet FS-2017-01 | version 1.1 | 4 April 2017

Traditionally, connections between mail servers have hardly been secured. STARTTLS is an extension to provide existing protocols with connection security. If you only use STARTTLS to secure connections between mail servers, this will protect against so-called passive attackers. An active attacker can easily undo the use of STARTTLS. The DANE protocol allows you to verifiably indicate that your server offers a secure connection.

The NCSC recommends enabling STARTTLS and DANE for all your organisation's incoming and outgoing email traffic.

Background

Traditionally, connections between mail servers have hardly been secured. The protocol for e-mail traffic, SMTP, is from 1982.[1] The designers of the protocol did not include a provision to require secure connections between mail servers. Therefore, many mail servers still allow unencrypted connections.

Target audience

Information security officers and administrators of email and DNS servers

The following organisations have contributed to this factsheet:

Atos, dmarcian, Dutch Standardisation Forum, municipality of 's-Hertogenbosch, KPN CISO, NLnet Labs, SIDN Labs, Sonnection, SSC-ICT

[1] The SMTP protocol was first standardised in RFC 821. The current version is RFC 5321: https://datatracker.ietf.org/doc/rfc5321/.

STARTTLS and DANE protect against different threats than OpenPGP and S/MIME

Multiple standards exist to send confidential emails securely. You can encrypt your emails end-to-end by using the standards OpenPGP or S/MIME. Thus, you encrypt the contents of the email itself. Someone listening in on the connection cannot read the contents of the email.

On the one hand, OpenPGP or S/MIME protects against threats against which STARTTLS and DANE do not protect. Even if an attacker penetrates the mail server, he cannot read the contents of the emails. On the other hand, STARTTLS and DANE protect *all* emails between mail servers that apply these standards. The user does not have to do anything to achieve this. Furthermore, STARTTLS and DANE protect the entire email, including metadata such as sender and subject. OpenPGP or S/MIME protects only the contents of the email.

STARTTLS is an extension to provide existing protocols with connection security. For SMTP, the use of STARTTLS is standardised in RFC 3207.[2]

What is the matter?

If you only use STARTTLS to secure connections between mail servers, this will protect against so-called ***passive*** attackers. STARTTLS prevents eavesdropping on a connection if an attacker only reads the contents of the connection and does not modify them. An attacker who behaves in such a manner is called a passive attacker. An ***active*** attacker, who does modify the traffic, can easily undo the use of STARTTLS. The attacker modifies the traffic in such a way, that the sending mail server thinks the receiving mail server does not support STARTTLS. He does this the other way around as well. This is commonly known as a STRIPTLS attack.

The DANE protocol allows you to verifiably indicate that your server offers, and prefers, a secure connection.[3] DANE is a protocol that uses DNS to offer information about connection security. This information is verifiable using DNSSEC.[4] A STRIPTLS attack is no longer possible if the sender and receiver of an email use DANE. If you use DANE, other mail servers can safely send messages to your mail server. Your organisation can also send email safely to organisations that use DANE for their mail servers.

The National Council Digital Government has decided in September 2016 to include STARTTLS and DANE for email traffic in the list of compulsory open standards.[5] Therefore, it is compulsory for Dutch government bodies to apply these standards when investing in email systems.

What could happen?

If you do not use STARTTLS and DANE to secure traffic from and to your mail servers, malicious parties can intercept the network traffic to and from your mail servers. This network traffic contains the contents of all emails that this mail server handles. An attacker needs to have access to the network traffic of your mail server to perform this attack. For a foreign intelligence agency or a criminal organisation, this is a realistic attack scenario.

It cannot be determined how many passive attacks occur on unencrypted traffic to and from mail servers. After all, the behaviour of the attacker is wholly invisible in this case. Such attacks are, however, highly profitable, as they provide attackers with much confidential information of the target organisation.

STRIPTLS attacks occur in practice. In 2015, researchers demonstrated that the STARTTLS protection to Google mail servers is stripped from more than twenty percent of all emails in seven countries. In certain cases, this percentage reached almost one hundred percent.[6] These emails were therefore sent unencrypted across the internet.

What does the NCSC recommend?

The NCSC recommends enabling STARTTLS and DANE for all the ***incoming*** email traffic of your organisation. In this way, every other organisation can safely communicate with your mail servers.

The NCSC further recommends enabling STARTTLS and DANE for all the ***outgoing*** email traffic of your organisation. Other organisations will also enable STARTTLS and DANE for their incoming email traffic. Your organisation can then communicate safely with other organisations that use STARTTLS and DANE on their incoming mail servers.

[2] See https://datatracker.ietf.org/doc/rfc3207/.
[3] This use of DANE is standarised in RFC 7672: https://datatracker.ietf.org/doc/rfc7672/.
[4] DNSSEC is standardised in RFC 4033 (https://datatracker.ietf.org/doc/rfc4033/), RFC 4034 (https://datatracker.ietf.org/doc/rfc4034/) and RFC 4035 (https://datatracker.ietf.org/doc/rfc4035/).

[5] See https://www.forumstandaardisatie.nl/nieuws/nationaal-beraad-verplicht-starttls-en-dane for more information.
[6] Source: Neither Snow Nor Rain Nor MITM...: An Empirical Analysis of Email Delivery Security, https://dl.acm.org/citation.cfm?id=2815695.

Making TLSA records

You make two TLSA records for every mail server. The first record, the '2 1 1' record, refers to the CA. The second record, the '3 1 1' record, refers to the certificate itself. In this way, minor mistakes in the configuration do not immediately lead to loss of availability of your email facility.[7]

Use the tool tlsa from the package hash-slinger to easily generate TLSA records.[8][9]

First, download the certificates of the server (server.pem) and the CA (CA.pem) to your workstation. In the examples, we assume that your mail server has FQDN mail.example.nl.

Creating a '2 1 1' record:

```
$ python tlsa --create --port 25 --usage 2
--selector 1 --certificate CA.pem
mail.example.nl
```

Creating a '3 1 1' record:

```
$ python tlsa --create --port 25 --usage 3
--selector 1 --certificate server.pem
mail.example.nl
```

STARTTLS and DANE on incoming email traffic

To implement STARTTLS and DANE on incoming email traffic, you first enable STARTTLS on every incoming mail server. Next, you publish TLSA records for these servers in the DNS zone of the servers. The DNS zones have to be protected with DNSSEC.

Make an inventory of the mail servers on which your organisation receives email. Include the mail servers that receive email from other external mail servers. These may include servers that are not under your control, such as servers of a spam filtering service. Include every server that is in the MX records of the domain name of your organisation. You may also secure internal email flows with STARTTLS and DANE. However, you can secure these email flows by other means as well, such as certificate pinning. Your organisation may already be using such measures.

Choose whether you offer STARTTLS based on a public certificate authority (CA) or a certificate authority of your own.[10] Use your own CA only if you have the knowledge and means to set it up and maintain it. Make sure that every mail server has its own certificate. Include the fully qualified domain name of the mail server in the certificate as Subject Alternative Name.

Enable STARTTLS on every mail server on your list. Configure STARTTLS based on the IT Security Guidelines for Transport Layer Security.[11] Use the certificate that you have created for this server. Configure the server in such a way that it sends the entire chain of certificates up to and including the CA.

Check whether the server is indeed reachable via STARTTLS. For example, use the email test on internet.nl to check this. Some firewalls are configured by default to strip STARTTLS from all incoming email traffic. If your server is not reachable via STARTTLS, change the network configuration the make the server reachable via STARTTLS.

For every mail server, publish information about the certificate and the CA in TLSA records in the DNS zone of the mail server. For example, if the mail server mail.example.nl handles the email of the domain example.org, you place the TLSA records in the DNS zone example.nl. See the frame 'Making TLSA records' for detailed instructions.

Make sure that DNSSEC is enabled, both on the DNS zone of the email domain and on the DNS zone that contains the TLSA records.[12] DNSSEC ensures that sending mail servers are able to verify the authenticity of information in TLSA records. Only with DNSSEC does publishing TLSA records have any effect.

Check regularly whether your configuration is correct and functional.[13] For example, use the mail test on internet.nl or the DANE SMTP tool of sys4.[14] Check whether your DNSSEC configuration is correct and functional as well. If you use DANE and STARTTLS for email, the availability of your email setup depends on DNSSEC.

[7] This method is based on the analysis from http://postfix.1071664.n5.nabble.com/WoSign-StartCom-CA-in-the-news-td86436.html#a86444 and follows the advice in the publication 'Trustworthy Email' of the NIST (https://www.nist.gov/node/1099976).
[8] See https://github.com/letoams/hash-slinger. Hash-slinger is also available in package managers of popular Linux distributions.
[9] You can also generate TLSA records with OpenSSL (an example is available in https://www.dnssec.nl/cases/tweeluik-dane-deel-ii-tlsa-records-voor-mail.html).

[10] In this scenario, the external use of an internal CA is permissible because the trust in the CA is derived from the information in DNS. Sending mail servers can verify this information using DNSSEC.
[11] See https://www.ncsc.nl/actueel/whitepapers/ict-beveiligingsrichtlijnen-voor-transport-layer-security-tls.html.
[12] For Dutch governments, the use of DNSSEC is compulsory since 2012 via the list of compulsory open standards: https://www.forumstandaardisatie.nl/standaard/dnssec.
[13] An overview of common mistakes in applying DANE and STARTTLS is available on https://dane.sys4.de/common_mistakes.
[14] See https://dane.sys4.de/.

Replacing certificates

Replace the certificate of a mail server if it expires soon, or if you suspect that an attacker was able to steal the private key.

First, generate the new certificate and have your own or the public CA sign it. Configure the mail server to use it. Then, change the TLSA record with type '3 1 1' of the mail server such that it refers to the new certificate.

Switching certificate authorities

The previous instructions for replacing a certificate assume that the new certificate was issued by the same certificate authority as the old certificate. In the period between replacing the old certificate and changing the TLSA record with type '3 1 1', validation occurs based on the '2 1 1' record.

If you wish to use a new certificate from another certificate authority, then first replace the TLSA record with type '2 1 1' by a TLSA record with type '2 1 1' that refers to the new certificate. Next, wait until the TTL (time to live) of the TLSA records has expired. Now, the old '2 1 1' record is no longer included in caches of DNS servers. Next, execute the procedure from the section 'Replacing certificates': generate a new certificate, have the new CA sign it, configure the mail server to use it and replace the '3 1 1' record.

STARTTLS and DANE on outgoing email traffic

To implement STARTTLS and DANE on outgoing email traffic, you enable DANE validation on every outgoing mail server of your organisation. This mail server needs to be able to perform DNSSEC validation.

Make an inventory of the mail servers that send email for your organisation. Include the mail servers that send email to other external mail servers. You may also secure internal email flows with STARTTLS and DANE. However, you can secure these email flows by other means as well, such as certificate pinning. Your organisation may already be using such measures.

For each mail server on your list, determine whether its mail server software supports DANE and STARTTLS for outgoing email. Consult the documentation of your mail server or ask your vendor.

If a mail server supports DANE and STARTTLS, enable them. Use the option to only perform DANE validation when TLSA records are available. This option may also be called 'opportunistic DANE validation'. Make sure that the mail server has a secure connection to a DNSSEC-validating recursive DNS name server. For example, run one locally on the mail server itself.

If a mail server does not support DANE but it does support STARTTLS, this mail server cannot automatically protect connections with external mail servers from active attackers. Ask the vendor of your mail server software when they will add support for DANE. As a temporary measure, set up a separate mail server as a relay for this mail server. On this relay, use mail server software that does support STARTTLS and DANE.[15] Use certificate pinning to secure the connection between the existing and the new server. Enable support for STARTTLS and DANE on the new server, according to the instructions from the previous paragraph.

Finally

If you wish to secure all your email traffic with DANE and STARTTLS, you have to implement it on your incoming and outgoing email traffic. However, you do not need to do this at the same time. For example, you may decide to start protecting all incoming traffic with DANE and STARTTLS right now, but to postpone enabling it for outgoing traffic. In this way, you are already reachable for all your contacts via a secure connection.

It is only useful to protect your incoming email traffic for an email domain with DANE and STARTTLS if you do that for all incoming mail servers of that email domain. Otherwise, an active attacker can block access to the protected mail servers, and thereby force an unencrypted connection. This does not hold in the same way for outgoing traffic. Every outgoing mail server that you protect with DANE and STARTTLS is an improvement.

In any case, enable STARTTLS for all your incoming and outgoing email traffic, even if you postpone implementing DANE. Against passive attackers, STARTTLS is itself an effective measure.

Monitor the validity of the certificates of your mail servers and the correctness of the TLSA records. There are two TLSA records for every mail server. At any given moment, one of these two has to be valid for your mail server to be reachable. Therefore, make sure that you notice problems before both TLSA records are no longer valid.

[15] At the time of publication, Postfix and Halon supported DANE and STARTTLS. The developers of Exim are working on supporting it.

Publication
National Cyber Security Centre (NCSC)
P.O. Box 117, 2501 CC The Hague
Turfmarkt 147, 2511 DP The Hague
+31 (70) 751 5555

More information
www.ncsc.nl
info@ncsc.nl
@ncsc_nl

FS-2017-01 | version 1.1 | 4 April 2017