Mitigating TLS compromise with ECDHE and SRP

Aron Wussler
Proton Technologies AG
aron@wussler.it
May 29, 2020

Abstract

The paper reviews an implementation of an additional encrypted tunnel within TLS to further secure and authenticate the traffic of personal information between ProtonMail's frontends and the backend, implementing its key exchange, symmetric packet encryption, and validation. Technologies such as Secure Remote Password (SRP) and the Elliptic Curves Diffie Hellman Ephemeral (ECDHE) exchange are used for the key exchange, verifying the public parameters through PGP signatures. The data is then transferred encrypted with AES-128-GCM.

This project is meant to integrate TLS security for high security data transfer, offering a flexible model that is easy to implement in the frontends by reusing part of the standard already existing in the PGP libraries.

1 Introduction

Transport Layer Security (TLS) is the protocol that secures most of our web traffic every day, while connecting to most of the websites we visit. It is often represented as a green lock next to the address of the webpage, this means the connection is encrypted and the content is loaded securely. The security of the connection is provided by a cipher suite, made up of an asymmetric encryption algorithm to exchange a key, and a symmetric one to transfer encrypted packets with the mediated key. The authenticity of the connection is guaranteed by a certificate, released by an authority that certifies the ownership of the domain.

This does not fully protect the user from an attack, several cases of TLS compromise are known [8], both malicious [7, 9, 10, 11] and consensual [6], predating on the naivety of a user or in corporate-issued devices. The most common situation is the inclusion of an additional (malicious) authority that is now allowed to issue rogue certificates and mark every connection as protected, defeating the purpose of TLS.

A high-security service, such as ProtonMail, is interested in adding a further layer of protection to the connection negotiating an encrypted tunnel within TLS, with its own certification system, key exchange, and symmetric encryption. The implementation of this tunnel is fully in the application layer, thus not dependent on the browser or the underlying operating system.

All the frontends connecting to ProtonMail's backend API will have to negotiate a session key as the first interaction, to be then used to encrypt every packet. While generally using tunnels within tunnels is frowned upon, because of the performance drawbacks compared to the slight security increase, this might be a useful niche application: the larger issue here is not the encryption scheme, but rather the certification system.

The key exchange leverages elliptic curves, further secured with a byproduct of the password exchange when logging in. To ensure the authenticity of this procedure, the public parameters of the exchange are signed with a PGP-key stored in ProtonMail's key verification system, based on the Merkle tree technology. This project, matched with proper source code transparency and verification, can easily prove the authenticity of the connection without compromises or intermediaries.

2 Design

2.1 Threat Model

TLS compromise is the main threat we are going to address with this countermeasure, it might seem an unlikely scenario but as we are going to see in the next paragraphs there is plenty of real life cases where this happened. This is mostly done on purpose, by large and medium corporate environments to surveil their users' traffic, with little care about certificate validation and best security practices, like not using always the strongest ciphers available.

In other, worse cases, it is compromised maliciously, by attackers and government entities with various techniques, most commonly by abusing the trust chain or exploiting implementation bugs, rather than cracking cryptography.

In this section these compromises are going to be analyzed, to examine security of TLS implementation until the latest 1.3 version.

2.1.1 TLS proxy

TLS proxies are used in many business to spy on the user's internet traffic to avoid data leak or to filter content, or in malicious settings they are also known as KCI (Key Compromise Impersonation) [8]. This malicious attack is mostly performed through social engineering and convincing an unknowledgeable user to install certificates.

The compromise consists in a MITM (Man In The Middle) attack: a root certificate is installed on the client, and when a connection to the server is done, a proxy, or TIA (TLS Intercept Application), intercepts it to act as a middleman, negotiating a session key both, with the client and the server, thus being able to snoop in the communication. The client displays no warning while this is happening because the chain of trust is guaranteed by the extra root certificate, hence not recognising the proxy's fake certificate as untrustworthy.
Up to TLS 1.2 [3] the TIA behaviour is the following:

- the TIA forwards the client’s ClientHello message as is to the server and inspects the resulting ServerHello and X.509 certificates;
- based on the name in the certificate, the TIA may:
  - “drop out” of the connection, i.e. allow the client and server to communicate directly;
  - interpose itself, answering the client with an alternative ServerHello and the server with an alternative ClientKeyExchange, negotiating different encrypted sessions with each party and forwarding so that it can see the plaintext of the connection, as shown in figure 1.

The attack can be performed also asynchronously, generating the server certificate based upon the ClientHello, see figure 2.

This makes the “drop out” of the encrypted session impossible, i.e. a TIA is unable to selectively choose which connections to monitor: once you proxy a connection, you have to proxy it until it’s done. Furthermore, the version 1.3 prevents downgrade attacks, disallowing a TIA that only supports TLS 1.2 to negotiate a TLS 1.3 connection, without triggering an illegal parameter error.

This is fully analyzed by Symantec [6] presenting a theoretical implementation model for TLS 1.3 traffic interception.

2.1.2 Abusing the CA infrastructure

The Certificate Authority infrastructure is built on trust, strong in theory, but in many cases CAs have been caught cheating in the wild [7]:

“...We are now witnessing occasions in which digital certificates are being spoofed or hijacked and then sold on the black market. Attackers have been successful in misleading trusted CAs, including Verisign, Comodo, and DigiNotar, into issuing fraudulent certificates for websites, such as Microsoft and Google. In September 2011, stories surfaced involving DigiNotar, a Dutch “default-trusted” root CA. Apparently, the CA may have been compromised maybe as early as 2009; 531 false certificates had been issued including certificates for domains such as Google, Facebook, Microsoft, Twitter, Mozilla, Mossad, MI6, and the CIA.”

These attacks raised concerns and new measures appeared, Google has been pushing for CT, Certificate Transparency, introducing transparency logs, in which all the issued certificates appear, to ensure that no rogue certificate is emitted. This is mostly a retrospective safety measure, losing the CA authority status is a good obstacle to anyone trying to abuse
the system, but does not guarantee safety a priori.

2.1.3 Hash collision

A practical exploit for MD5 [9] has already been executed in 2005, raising many concerns. By now both MD5 and SHA-1 algorithms have been deprecated, many browsers and applications will display warnings or refuse connection when presented with these hashes.

The practical solution found by the researchers was to alter specific bits of the hard to factor moduli to present identically signed certificates and forge a certificate that shared the same signature, hence being verified by a CA, but holding different private keys, allowing an eavesdropper to intercept traffic.

2.1.4 Weak cipher suites

Many TLS cipher suites have been compromised [10]: BEAST (on CBC encryption in TLS 1.0), Lucky13 (on CBC padding in TLS 1.2), POODLE (on CBC padding in SSL 3.0), RC4 NOMORE (on keystream biases in RC4), FREAK (on export-grade RSAkeys), Logjam (on export-grade Diffie-Hellman groups), and DROWN (on RSA-PKCS#1v1.5 encryption in SSL 2.0).

Leaving these chiper enabled in a webservor or client is bad practice, but widely done not to drop connection for older clients. Many of this scenarios are vulnerable to downgrade attacks by exploiting protocol flaws (e.g. Logjam, SLOTH), and flawed cipher suites should be disabled entirely.

In 2016 researchers were able to exploit 64-bit block cipher [10] intercepting few hundred gigabytes of data and using a birthday paradox attack on it.

2.1.5 Compression attacks

Starting with CRIME, Compression Ratio Info-leak Made Easy, in 2012 a series of attack against TLS compression have been done. An attacker that will be able to induce a client to perform some specifically crafted requests will be able to use the size of the response as indicator of the content of the request. This is especially true with CSRF tokens as demonstrated by BREACH attack [11] that will reveal secrets repeated in the body in less than a minute.

This types of attacks do not rely on crypto itself, as they can be used against stream ciphers as well as block ciphers, rather on implementation itself, thus being much more insidious.

2.2 Structure

The tunnel is designed to protect the data since the first API call because username, registration, and password recovery are critically sensitive information that if the protection started with SRP would not be encrypted. Therefore, an ECDHE exchange is performed on landing to mediate an encryption key. In detail this exchange goes as follows:

1. the client requests the server key $Q$ with a GET request at /tunnel/key;
2. the server replies with a signed $Q$ point on the curve25519 and the server’s signing PGP key;
3. the client verifies the PGP key against a Merkle tree for key transparency;
4. the client generates the public parameter $V$ and the secret $Z$, submitting $V$ at POST /tunnel/key;
5. the server calculates the secret $Z$, and returns a newly generated session UID;
6. client and server hash $Z$ with the PGP key fingerprint to obtain a session key.

From now on, all the traffic is going to have only the headers necessary for transmission and the UID in clear to allow the decryption. The private signing PGP key is meant to be kept airgapped, so that a rogue server will not allow an attacker to know the private signing key, that can be used to re-sign a new $Q$. This procedure has the drawback that it requires an administrator to rotate them manually.

When the log-in happens SRP generates a secret $K$ that can be used for the tunnel:

1. the client sends the server the username with a POST request at /auth/info;
2. the server replies with an ephemeral challenge $S$, a modulus $M$, and a salt;
3. the client computes the ephemeral parameter $C$, the secret $K$, and $P_C = \text{Hash}(C, S, K)$ as authentication proof, then sends $C$ and $P_C$ at POST /auth;
4. the server then calculates $K$ using $C$, verifies $P_C$, generates $P_S$ to prove its identity to the client, and sends $P_S$ back;
5. the secret $K$ is finally hashed by both parties and used as session key;

As it can be observed the secret $K$ is generated with a personalised challenge for every user, using their own password, hence making it very difficult to crack this key on an extended scale. For every user and connection the discrete logarithm problem must be solved, without giving a single failure point as the factorization of a public key.

It is worth underlining that an attacker could save the traffic for a long time to crack until it is able to find the password to crack it all at once, but the secret $K$ cannot be derived just knowing the password, hence preventing this scenario. In that case a new session can be created, but no information can be derived onto a past or existing one.

The cipher suite Curve25519 with AES-128 was selected because of its robustness and convenience: the algorithms are already present into the OpenPGP standard [3], making implementation on clients easier, no additional library is necessary.

To make the tunnel more robust and to avoid birthday attacks the keys are refreshed after a set amount of time, enforced by the server with an error. The client is then allowed to use the old key just to refresh the session. Since this exchange is encrypted the only information leaked is the timing and size of the packets, making it very difficult for an attacker to guess when is this happening.
Figure 4: MD5 colliding certificates' public key moduli [9]

Figure 5 visualises this mechanism, in particular the user actions and the corresponding key exchanges. It is important to understand that since exchanges are performed within a tunnel, to fully compromise the transferred information both exchanges must be broken, as even cracking the shared ECDHE $Q$ public parameter will not provide information after the SRP exchange, keeping the users’ keys safe.

Once exchanges are performed, AES-128-GCM is used as symmetric encryption, as it is authenticated and fast, taking advantage of hardware acceleration on the server and webcrypto on the compatible clients.

In figure 6 is represented the packet encryption and decryption mechanism: a plain RFC 791 packet is encrypted before leaving the client. The necessary headers are kept unencrypted, and the rest is encrypted with the same structure provided by the standard: CRLF to separate each header and two times CRLF to separate the headers from the body. In this way all types of packet can be easily transmitted, whether they are plaintext, JSON, or binary data.

To prevent replay attacks the packet is equipped with a nonce and a timestamp, a packet is allowed 120s to get to the server, after which is no longer considered valid. In this timeframe the packets must always have a unique nonce, so that an attacker sending twice the same packet will have the second bounced before it is processed.

In table 1 the precise structure of a packet is outlined. The encrypted part consists of the first 16 bytes of initialisation vector (IV), followed by the AES-128-GCM encrypted data and its 12 byte authentication tag. In this way we ensure that data cannot be modified by altering the IV or by tampering with the ciphertext.

The system is built to handle cookies both in the encrypted and unencrypted subsection, as the webclient uses secure cookies that can not be accessed from the javascript application, hence making it impossible to encrypt them. Since cookies store only the access and refresh tokens, that are associated with an encrypted session, they are useless without the encryption key, so can be transmitted in clear without any security trade-off.

Client and server, when decrypting packets will join the headers to re-create the original packet, in case of conflicting headers the encrypted ones will prevail. This applies to cookies too, since they are in practice transmitted as a header. A limitation of this method is the unability to send both, encrypted and unencrypted cookies, for instance the aforementioned case of the secure cookies.

### 3 Implementation

#### 3.1 Backend implementation

This project is composed of several parts, an API controller, a Session model, and a middleware. The controller provides a three RESTful routes:

- **GET /tunnel/key** to request the signed public parameter $Q$, a point on the curve25519, to perform the ECDHE;
- **POST /tunnel/key** to submit the ECDHE exchange and issue a new tunnel session, identified by an UID;
- **PUT /tunnel/key** to submit an ECDHE exchange to refresh the tunnel when it is expired.

This API relies on the Session model to store the derived encryption key and expiration, identified by the unique identifier UID. When a client connects through the tunnel an encrypted session is created, that cannot in any case be converted to an unencrypted one. A session is uniquely tied with an oauth AccessToken and RefreshToken through the UID, this means

| Unencrypted (Headers) | Necessary headers for transmission and decryption |
|-----------------------|--------------------------------------------------|
| Timestamp, nonce, session, cookies* | Initialization Vector (IV) |

| Encrypted (Body) | Payload | Body of request or response |
|------------------|---------|-----------------------------|
| Initialisation Vector (IV) | Timestamp, nonce, session, cookies* | Body of request or response |

| Table 1: Tunnel packet structure |
|---------------------------------|

| Unencrypted (Headers) | Necessary headers for transmission and decryption |
|-----------------------|--------------------------------------------------|
| Timestamp, nonce, session, cookies* | Initialization Vector (IV) |

| Encrypted (Body) | Payload | Body of request or response |
|------------------|---------|-----------------------------|
| Initialisation Vector (IV) | Timestamp, nonce, session, cookies* | Body of request or response |
that if the encryption key is lost then the user must login again and a new session must be established.

Once an encrypt session is established all requests to the API will be routed through POST /tunnel/data, to be then handled by a middleware, schematically represented in figure 7. It is executed as an outer shell of the application: it intercepts the requests, decrypts and verifies the validity of the packet, and if any of the parameters is wrong it will return an error to the client. If the server is able to decrypt the request the response or error is encrypted, otherwise it is unencrypted.

After verifying the previously discussed security features, such as nonce and timestamp, the middleware creates a new request with the unencrypted route, parameters and payload, ready to be submitted to the router. This component is run as an early middleware, allowing it to intercept the request almost before any parsing is done. The API, being written in PHP, relies on some pre-processing that unfortunately is not available within the encrypted tunnel and is to be done manually, specifically the superglobals: $GET, $POST, $COOKIES, $SERVER, and $FILES.

Encrypted headers, once decoded, are added to the new request merging them with the $SERVER superglobal. Since cookies are also a header, they are specifically parsed by a class according to RFC 6265 [4], subsection 5.4.

The API mostly handles JSON encoded requests, with some exceptions, such as attachment submission or file upload. The discriminant factor is the Content-Type header, when it is either multipart/form-data or application/x-www-form-urlencoded the request is additionally parsed by a class built upon RFC 2388 [1]. Many edge cases and the necessity to parse data coming form a multitude of different clients and browsers made this task particularly intricated.

3.2 Frontend implementation

3.2.1 Client threats

The tunnel has been designed to address the threat model, given that the front end code is not flawed. On application clients, like the Android or iOS app, as well as the bridge and the VPN apps this is easily addressable, while for the web-client extra safety measures must be implemented, to verify the JS app integrity. This project is being developed parallel to a code integrity verification system, that will provide trust on the application layer.

Key Verification To verify the signature on the public parameter $Q$ of the ECDHE exchange a public signing key is available in ProtonMail’s key verification system. This technology is based upon COINKS [12], an end-user key verification service capable of integration in end-to-end encrypted communication systems. The clients will have to first retrieve the key from the public key API, then verify it against the Merkle tree as described in the paper, to then verify the signature.
3.2.2 Migration

Clients must be allowed to communicate in both the encrypted and unencrypted form; since it will require time to implement this feature the tunnel will not be enforced on all connections. If at the beginning of a session an ECDHE key is negotiated between the two parties then the connection will be stored as encrypted and it will not be possible to communicate without encryption using that specific identifier, as it could be used as a downgrade attack.

Sessions are in fact divided in encrypted and unencrypted and it won’t be allowed to switch between the two states. Disabling encryption on a session can be overridden by the server, for instance when testing, since the developer tools in any browser will just show unintelligible blobs of data being sent to POST /tunnel/data, and using any API development tool will turn out to be overly complicated.

3.3 Testing

To test the API implementation several approaches have been followed: unit tests for the single modules, functional tests for the API and middleware, end-to-end tests for a bigger-picture testing, and a proxy for manual testing.

Unit and functional tests have been developed to work on Continuous Integration (CI), and be automatically performed, while end-to-end must be run with a fully functioning HTTP server to test the middleware’s interaction with the requests.

All of these tests are designed to test several API functionalities, such as key exchanges and renegotiation, and the middleware functionalities, from the basic encryption and decryption to the edge cases of encrypted cookies and multipart requests.

3.3.1 Test proxy

To perform manual testing of the tunnel a NodeJS proxy has been developed, nicknamed NodeProxy, that provides a compatibility interface between the existing clients and the API with the tunnel enabled.

To mediate this interface the test proxy:

1. performs the key exchange using the OpenPGPJS library, directly accessing the newly-inserted genPublicEphemeralKey function;
2. initiates a session and displays the current session UID and encryption key, taking care of encryption and decryption of all traffic;
3. on login intercepts the SRP params and performs a MITM attack and updates the encryption key;
4. when key refresh is necessary it locks and queues all the requests in an await status, to perform a new handshake.

This proxy is completely transparent to the clients, it is needed just to alter the API address to match the proxy’s.

In order to perform SRP transparently this tool is requires the user’s password and features a full implementation of an inline MITM attack. The proxy:

1. detects a log-in attempt by listening to the API calls, when a call to POST /auth/info a verifier and a challenge are generated, emulating an SRP server;
2. forwards the signed modulus in the server’s response, and alters the ephemeral challenge;
3. detects a call to POST /auth and generates a client ephemeral and proof;
4. responds to the client with a newly-generated server proof and confirms the login.

This attack, illustrated in figure 8, is possible only by knowing the user’s password, and required writing a JS implementation of an SRP server, previously not available.

4 Conclusions

The aim of this project is to guarantee the security and authenticity of all the information shared between the client and the server by encrypting every request to the server, before
transmitting it over TLS. This is achieved by studying and addressing TLS’s main weaknesses, thus significantly raising the complexity of an attack.

In most cases TLS is being systematically intercepted by proxies, that would be prevented to access and alter the transmitted information, thus only being capable to block the user from connecting ProtonMail’s services without being able to tamper with the data. Malicious TLS compromise is very alarming because it is a critical issue with a very little footprint, as it is usually a relatively elaborate attack, that requires a targeted intent. In this case cracking the second layer of encryption renders the attack much more difficult, either by altering the client code or by cracking another set of encryption keys.

The tunnel secures the packets by encrypting them with a symmetric cipher, keeping public only the session identifiers necessary to fetch the correct decryption key. In order to mediate a session key, two different key exchanges are used: the first using elliptic curves immediately on landing, the second taking advantage of the residual secret from the password exchange protocol. Login is therefore a critical security step for the tunnel: assuming an attacker does neither know the password verifier, nor the password itself, the SRP exchange is immune to MITM attacks. In fact, the tunnel not only protects the transmitted data, but ensures full authenticity of the connection, guaranteeing that every non-compromised client implementation is communicating exclusively with ProtonMail’s authorized servers.

Nevertheless, this does not mean that TLS is broken, and this implementation does not supply a replacement to it. The tunnel can indeed be used to transfer data over an insecure channel, but it is designed to co-operate with TLS and remedy to the most critical aspects of it. Moreover, this project is not meant to be an implementation standard as it does not fit most use modes of TLS. In fact, the project derives from a further need of data protection for high-confidence applications and is tailored onto ProtonMail’s specific case.

Acknowledgments

I would like to thank Francisco Vial for the help reviewing this paper.

References

[1] L. Masinter, RFC 2388: Returning Values from Forms: multipart/form-data, Network Working Group, August 1998 https://www.ietf.org/rfc/rfc2388.txt
[2] W. Koch, B. Carlson, R.H. Tse, D.A. Atkins, D.K. Gillmor, RFC 4880bis draft 09: OpenPGP Message Format, Network Working Group, March 2020 https://tools.ietf.org/html/draft-ietf-openpgp-rfc4880bis-09
[3] T. Dierks, E. Rescorla, RFC 5246: The Transport Layer Security (TLS) Protocol Version 1.2, Network Working Group, August 2008 https://www.ietf.org/rfc/rfc5246.txt
[4] A. Barth, RFC 6265: HTTP State Management Mecha-
Figure 8: SRP MITM attack, inline, as implemented in NodeProxy. $M$ is the fixed Modulus, $g$ the generator, $p$ is the user’s password, $v_s$ is the server’s password verifier.