Post-Quantum Authentication in TLS 1.3: A Performance Study

Dimitrios Sikeridis*, Panos Kampanakis†, Michael Devetsikiotis*

* Dept. of Electrical and Computer Engineering, The University of New Mexico, Albuquerque, NM, USA
† Security & Trust Organization, Cisco Systems, USA
{dsike, mdevets}@unm.edu, panosk@cisco.com

Abstract—The potential development of large-scale quantum computers is raising concerns among IT and security research professionals due to their ability to solve (elliptic curve) discrete logarithm and integer factorization problems in polynomial time. All currently used public key algorithms would be deemed insecure in a post-quantum (PQ) setting. In response, the National Institute of Standards and Technology (NIST) has initiated a process to standardize quantum-resistant crypto algorithms, focusing primarily on their security guarantees. Since PQ algorithms present significant differences over classical ones, their overall evaluation should not be performed out-of-context. This work presents a detailed performance evaluation of the NIST signature algorithm candidates and investigates the imposed latency on TLS 1.3 connection establishment under realistic network conditions. In addition, we investigate their impact on TLS session throughput and analyze the trade-off between lengthy PQ signatures and computationally heavy PQ cryptographic operations. Our results demonstrate that the adoption of at least two PQ signature algorithms would be viable with little additional overhead over current signature algorithms. Also, we argue that many NIST PQ candidates can effectively be used for less time-sensitive applications, and provide an in-depth discussion on the integration of PQ authentication in encrypted tunneling protocols, along with the related challenges, improvements, and alternatives. Finally, we propose and evaluate the combination of different PQ signature algorithms across the same certificate chain in TLS. Results show a reduction of the TLS handshake time and a significant increase of a server’s TLS tunnel connection rate over using a single PQ signature scheme.

I. INTRODUCTION

Digital communications have completely penetrated everyday life and the physical world as enablers of numerous critical services including telemedicine, online banking, massive e-commerce, machine-to-machine automation, mobile and cloud computing. In this reality, public-key cryptography is ubiquitous in almost all cryptographic protocols, such as TLS which builds encrypted tunnels between digital entities. Studies suggest that over 60% of Internet connections are implemented over the TLS-based secure HTTPS protocol [20], [64]. TLS adoption is expected to keep increasing as users and client vendors strive for ubiquitous encryption and privacy [51].

Apart from connection integrity and confidentiality, TLS provides authentication usually with the use of X.509 certificates [45]. Such certificates are issued by trusted third-parties called Certificate Authorities (CAs). Endpoints verify the communicating peer’s identity and public key (PK) contained inside his certificate by leveraging a chain of certificates that is rooted to a pre-trusted root CA. The two most popular digital signature algorithms used in certificates today are the Elliptic Curve Digital Signature (ECDSA) and Rivest-Shamir-Adleman (RSA). Their security guarantees rely on the hardness of the elliptic curve discrete logarithm (ECDL) and integer factorization (IF) problems respectively.

While the security of the aforementioned schemes cannot be practically challenged by conventional computer systems, this would not be the case in a post-quantum world where a large scale quantum computer has become a reality [62]. Shor’s quantum algorithm [76], [87], assuming a practical quantum computer (QC) was available, would solve ECDL and IF problems in polynomial time which would render ECDSA and RSA insecure. In this scenario, a QC-equipped attacker would be able to impersonate signers that use these algorithms. Thus, encrypted tunnel (e.g., TLS, IKEv2) authentication, PK infrastructure (PKI), CAs, and software signing would be broken.

To address the issue, the cryptographic community has been researching quantum-resistant public key algorithms for some time, while the US NIST started a public project to standardize quantum-resistant public key encapsulation and signature algorithms. Similarly, ETSI has formed a Quantum-Safe Working Group [32] that aims to make assessments and recommendations on the various proposals from industry and academia regarding real-world deployments of quantum-safe cryptography. Moreover, the IETF has seen multiple proposals that attempt to introduce and investigate PQ algorithms in protocols like TLS and IKE [33], [40], [68], [91], [93].

At the moment of this writing, NIST’s evaluation process has moved from Round 1 to Round 2 where 26 PQ algorithms were chosen with security guarantees being the primary criterion, while performance was treated as a future goal [2]. Evidently, the actual integration of these algorithms into existing protocols (e.g., TLS, IKEv2, SSH) and their coexistence with today’s Internet infrastructure present challenges that pertain to (a) additional latency due to their heavy operations, (b) communication overhead from the increased public key and signature sizes, and (c) optimal use of existing hardware towards faster implementations. These gaps are
actively being studied by research teams in the industry with NXP Semiconductors, Microsoft and Queensland University of Technology [15], Amazon [18], Cloudflare and Google focusing on the impact of key exchange mechanisms (KEMs) on TLS [52], [54]–[56].

These efforts mostly focus on TLS PQ key exchange as confidentiality is considered more urgent. Since a QC-equipped attacker would be able to decrypt stored communications retroactively, ensuring quantum-resistant encryption for critical data is a priority. This is not the case with authentication as digital entity impersonation cannot happen retroactively. However, there are numerous incentives that drive the early study of PQ authentication in today’s protocols. First, PKI refresh cycles are traditionally long and migration to new primitives can take years. A case in point is ECDSA. Even though it was standardized in 2005 [4] offering clear performance advantages over RSA [39], its adoption was still not broad a decade later. Finally, the computational performance of PQ algorithms, along with the fact that the sizes of the resulting PQ certificates are significantly larger, will definitely impact the TLS handshake by worsening user experience and connection performance. Thus, it is important to investigate and identify promising PQ signature candidates specifically for utilization in TLS.

In this paper, we study the overhead introduced by PQ certificates in the establishment of TLS 1.3 tunnels. Our goal is to identify PQ signature candidates that could be employed in TLS without major protocol updates, measure their performance in real-world deployments and contribute to the overall discussion about their use in encrypted tunnels. The key contributions of our work are summarized as follows:

(i) We analyze the candidate signature algorithms from NIST’s Post-Quantum Cryptography Project and compare them in terms of performance, security level claims and key/signature sizes. In addition, we integrate software implementations of these schemes in X.509 certificates for TLS 1.3 authentication in the OQS OpenSSL library [74].

(ii) We conduct large-scale measurements to investigate the PQ authentication algorithm impact on TLS 1.3 handshake establishment in realistic network conditions. Moreover, we investigate the impact that larger PQ certificate chains or slower sign/verify operations have on the throughput of a server performing PQ authentication in HTTPS.

(iii) We demonstrate the viability of adopting two PQ candidate signature algorithms, Dilithium, and Falcon, for time-sensitive applications over TLS and we argue that Falcon is more suitable for the web if floating point hardware is available at the server. We argue that less time-sensitive applications can use a larger set of PQ candidate algorithms.

(iv) We propose and study the use of different PQ signature scheme combinations in the same certificate chain to improve the performance of PQ authentication in TLS which, to the best of our knowledge, has not been examined before.

(v) We provide an in-depth discussion of PQ authentication challenges for encrypted tunnelling protocols, alternatives, and present insights towards optimizing future deployments.

Note that we do not evaluate the PQ security claims or security proofs of the PQ algorithms. These will be assessed in the NIST standardization process. We also do not perform an exhaustive benchmark of all the available parameter sets of the signature algorithms under consideration by NIST. Relying on preliminary findings and individual algorithm claims, we have explicitly chosen to focus on a subset of algorithms and parameter sets that would seamlessly fit in TLS. To the best of our knowledge, this is the first work that assesses the performance of PQ certificates in TLS 1.3 by considering realistic network conditions.

The rest of the paper is organized as follows: Section II presents background on X.509 PKI and TLS 1.3. In Section III, the different PQ candidate signature families and algorithms are presented, while Section IV details the integration of PQ authentication into TLS 1.3. Section V presents the experimental procedure, results and their analysis. Section VI summarizes related work. Finally, Section VII discusses the general implications and potential solutions of integrating new PQ signatures in encrypted tunnels, while Section VIII concludes this paper.

II. BACKGROUND

This section presents an overview of the TLS 1.3 handshake protocol, along with a summary of the X.509 PKI currently used in TLS.

A. X.509 Certificates and PKI

A digital entity’s (e.g., a server) identity is bound to its public key via a digital certificate. Since X.509 is the most common PKI standard adopted by IETF protocols, X.509 certificates play an important role in digital authentication for various protocols (e.g., TLS, SSH, IKEv2). X.509 certificates are defined by RFCs 5280 [22] and 6818 [94], and their sizes usually vary between 0.5 and 1.5KB.

A PKI infrastructure consists of various parts. A Certificate Authority (CA) issues an entity’s certificate that assures the entity’s identity and public key tie to that identity. The identity is included in the Subject field of the certificate, while the entity’s public key is stored in the Subject Public Key Information along with the algorithm used by the issuer to create the signature. A certificate contains a specific validity period and extensions included by CAs to enable additional functionality. The certificate is signed by the CA’s private key using the specified signature algorithm and the signature is added to the certificate’s Signature field.

Certificates are exchanged between entities during a session setup in order for each party to verify the peer’s identity. At the top of the X.509 PKI there are trusted CAs that self-sign their own certificates known as root CA certificates. Normally a root CA issues certificates for intermediate CAs (ICAs). Following that, the root CA is kept offline for security purposes. An ICA can further issue certificates for other ICAs that in turn sign leaf certificates in the PKI. This process results to the creation of certificate chains of trust that usually consist of two to four certificates but can be arbitrarily long. A leaf certificate is validated by an endpoint if (a) the endpoint trusts the chain’s root CA, and (b) all the signatures from the leaf to the root certificate of the chain are verified by using the public key of the issuer.
B. TLS 1.3 Encrypted Tunnels

The TLS protocol design provides endpoint authentication and establishes encrypted communication tunnels between them. The aim is to ensure data integrity, confidentiality, and authenticity. In August 2018, the latest version of the protocol — TLS 1.3 — was published as RFC 8446 [78]. TLS 1.3 offers significant improvements over its predecessor TLS 1.2 that include elimination of insecure or obsolete features, complete encryption of the handshake, and reduced handshake latency by eliminating a round-trip.

Below we present the full TLS 1.3 handshake, and a summary of the messages exchanged between a client that initiates a connection with a remote server: Initially, the client calculates an ephemeral public/private key pair for the key exchange and sends to the server a ClientHello message that contains a random nonce, the protocol versions and cipher suites that the client supports, and possible extensions which include pre-shared keys, a list of supported signature algorithms (signature_algorithms extension), or a list of supported signature algorithms specifically for certificates (signature_algorithms_cert extension).

In turn, the server calculates its own ephemeral key pairs, determines the desired cryptographic parameters and responds with a ServerHello message that contains the server’s nonce, a public key for the key exchange, the preferred cipher suite and optionally key_share and pre_shared_key extensions. At this point, by combining the ClientHello and ServerHello messages the two entities generate a shared key and the connection gets encrypted. The server continues by sending a ServerChangeCipherSpec message for TLS 1.2 compatibility purposes (“middlebox compatibility mode”). This is followed by the server sending ServerEncryptedExtensions extensions, and optionally a ServerCertificateRequest message if client certificate authentication is required. The encrypted extensions include the server’s certificate and certificate chain for authentication (ServerCertificate), and a ServerCertificateVerify message that contains a signature over the handshake. Finally, the server ends its part of the handshake with a ServerFinished message that contains a Message Authentication Code (MAC), namely verification data generated by a hash of all the messages exchanged so far.

Following that, and again for compatibility purposes, the client responds with a ClientChangeCipherSpec message to the server. If client authentication is requested the client then sends its public key certificate (a ClientCertificate message) and optionally a ClientCertificateVerify. The TLS 1.3 handshake finishes with the client ClientFinished message. The complete TLS 1.3 handshake and its variations are detailed in RFC 8446 [78], while an illustrated byte-per-byte description can be found in [29].

III. Post-Quantum Candidate Signature Schemes

In this section we present the different families of PQ signature algorithms, and document details regarding the specific schemes and parameter sets used in this study.

A. Quantum-Resistant Families of Problems

Since currently used public-key cryptography schemes would be threatened by large-scale quantum computers, research for alternatives — namely post-quantum schemes — able to resist QC attacks has been surging for the last decade. The goal has been to identify suitable problems or subproblems of NP-hardness that are not solvable in polynomial time by quantum algorithms [16].

Hash-based Cryptography This family of PQ signature algorithms relies on Merkle trees and few or one-time-signature (FTS/OTS) used with secure cryptographic hash functions. Important security requirements for these functions include collision and preimage resistance. The first scheme in the family, Merkle signature scheme (MSS), was presented in the late 1970s [16]. The use of hash functions with Merkle trees and FTS/OTS for generating signatures is considered mature, well-understood, and significantly reliable. Hash-Based Signature (HBS) schemes generate keypairs for the FTS/OTS. The FTS/OTS signs a message and the corresponding public key becomes a leaf in the Merkle tree. The resulting Merkle tree root is the public key. Currently, the most mature schemes of this family are the stateful LMS [60] and XMSS [43], and stateless SPHNICS+, one of the NIST signature candidates [5]. A stateful HBS relies on an OTS and a signer needs to ensure that the OTS private key is never reused. This state management requirement is considered an important disadvantage. While stateless SPHNICS+ alleviates this issue, it also leads to an increase in signature size to ~40-60 KB and slower performance.

Lattice-based Cryptography Another family of hard problems rely on lattices [65], [70]. A lattice is the set of all integer linear combinations of linearly independent vectors in real n-space \( \mathbb{R}^n \). There are many lattice-related NP-hard problems used for cryptographic purposes including the shortest vector problem (SVP) (i.e., find a shortest vector in the Euclidean norm), the closest vector problem (CVP) (i.e., find a lattice vector that minimizes the distance from another target lattice), and the lattice basis reduction problem. The lattice problems NIST PQ signature candidates depend on learning with errors (LWE) [77], ring learning with errors (RLWE) [59], module learning with rounding or errors (MLWR or MLWE) and NTRU [41]. Some of these may be reducible to these NP-hard lattice problems. The NTRU and LWE families have been studied more extensively than others. Lattice-based algorithms are promising quantum-resistant solutions with relatively efficient implementations and strong security properties. Among the NIST PQ signature candidates the list of lattice-based schemes includes Dilithium [30], qTesla [5], and Falcon [39].

Multivariate Cryptography Another family of problems that are used by some NIST’s signature algorithm candidates is related to solving multivariate quadratic equations over finite fields which is an NP-hard problem. The system’s hardness to solve depends mainly on the degree, the number of variables, and the underlying finite field’s size. Multivariate PQ schemes often lead to excessive key or signature sizes. Thus, recent research has focused on reducing keys sizes for these schemes [16]. The NIST multivariate-based signature candidates are Rainbow [27], MQDSS [21], LUOV [12], and GeMSS [19].
use in TLS, and by taking into account their performance and sets for each algorithm in our study was done specifically for bits of security respectively. The choice of specific parameter to 128-bit security, while Levels 3 and 5 offer 192 and 256 in our study, we focus on Level 1, 3 and 5. Level 1 corresponds on the bits of security offered by the algorithm parameter set.

### B. PQ Signature Algorithms and Parameter Sets Studied

In their PQ Project, NIST defined five security levels based on the bits of security offered by the algorithm parameter set. In our study, we focus on Level 1, 3 and 5. Level 1 corresponds to 128-bit security, while Levels 3 and 5 offer 192 and 256 bits of security respectively. The choice of specific parameter sets for each algorithm in our study was done specifically for use in TLS, and by taking into account their performance and signature/key size outputs.

**SPHINCS\(^+\) [5]:** The SPHINCS\(^+\) signature algorithm specification defines 36 different parameter sets, that utilize different hash functions including Haraka, SHA256, and SHAKE256. For our experiments, we chose to integrate and evaluate the SHA256-128f-simple parameter set that corresponds to NIST’s Level 1 because it was the most efficient one integrated in PQClean [75].

**Dilithium [30], [31]:** In our study, we integrated and evaluated Dilithium II which was presented at the Second PQC mailing list thread, G. Zaverucha & D. Kales offered a new attack which the MQDSS team acknowledged and will address in the future with a new parameter set that will probably worsen our measured performance.

**Falcon 1024 [34]:** Regarding the Falcon signature scheme, its Round 2 submission describes two parameter sets, Falcon 512 that provides NIST’s Level 1 security, and Falcon 1024 that corresponds to Level 5. In our experiments we integrated and evaluated the version provided by the Falcon team in liboqs [73], [90]. That version did not include the floating point hardware optimizations that improve Falcon’s signing performance by \( \sim 20\) times [72].

**qTesla [3]:** Initially, the heuristic version of the qTesla I parameter set (NIST Level 1) tested as a very promising TLS candidate. However, at the time of this writing, an update on qTesla’s Round 2 submission eliminated all Round 2 heuristic parameter sets. In their place the qTesla team kept the provably-secure parameter sets that result in significantly bigger signature (2.5KB) and public key (\( \sim 15\)KB) which would double the handshake time in TLS. Therefore, we did not evaluate qTesla further.

**Picnic [38]:** In this study, we examine the Round 1 Picnic L1FS parameter set which existed in the OQS OpenSSL [74] and corresponds to NIST’s Level 1. The Round 2 submission preserves all Round 1 parameter sets, and adds another three that are referred to as Picnic2. This new parameter set reduces the size of the produced signature by making the whole scheme significantly slower while the signature is still over 10KB [37], § 11.3, § 11.4. Thus, although our evaluation of Picnic is slightly optimistic, it still does not perform well in TLS.

**MQDSS [21]:** The MQDSS specification provides multiple parameter sets, and specifically recommends two of them. MQDSS31-48 corresponds to NIST Levels 1-2. In our study, we integrated and evaluated the MQDSS31-48 implementation from liboqs [73], [90]. Note that, in a recent NIST mailing list thread, G. Zaverucha & D. Kales offered a new attack which the MQDSS team acknowledged and will address in the future with a new parameter set that will probably worsen our measured performance.

**Rainbow [27]:** The specification of Rainbow suggests three different variants, and parameter sets. By examining the cryptographic operations’ performance, along with the public keys of each alternative, we chose to test the Rainbow Ia - Cyclic parameter set integrated in the PQClean project [75] because it was the most promising one for TLS. Rainbow Ia offers Level 1 security.

Table 1 summarizes the conventional, and PQ signature algorithms examined in this study and presents the sizes of public keys, private keys, and signatures.

<table>
<thead>
<tr>
<th>Signature Algorithm</th>
<th>Specification</th>
<th>Hard Problem</th>
<th>Public Key Size (Bytes)</th>
<th>Private Key Size (Bytes)</th>
<th>Signature Size (Bytes)</th>
<th>Claimed Classical Security Level</th>
<th>Claimed PQ Security Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA 3072</td>
<td>[61]</td>
<td>Integer Factorization</td>
<td>387</td>
<td>384</td>
<td>384</td>
<td>128 bits</td>
<td>( \sim 0) bits</td>
</tr>
<tr>
<td>ECDSA 384</td>
<td>[4]</td>
<td>EC Discrete Logarithm</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>192 bits</td>
<td>( \sim 0) bits</td>
</tr>
<tr>
<td>Dilithium II [30], [31]</td>
<td>Module Learning with Errors</td>
<td>1184</td>
<td>2800</td>
<td>2044</td>
<td>100 bits</td>
<td>91 bits</td>
<td>100 bits</td>
</tr>
<tr>
<td>Falcon 512</td>
<td>[34]</td>
<td>NTRU</td>
<td>897</td>
<td>1281</td>
<td>690</td>
<td>114 bits</td>
<td>103 bits</td>
</tr>
<tr>
<td>Picnic L1FS</td>
<td>[38]</td>
<td>Zero-Knowledge Proofs</td>
<td>33</td>
<td>49</td>
<td>34036</td>
<td>128 bits</td>
<td>64 bits</td>
</tr>
<tr>
<td>SPHINCS(^+) SHA256-128f-simple</td>
<td>[5]</td>
<td>Hash-Based</td>
<td>32</td>
<td>64</td>
<td>16976</td>
<td>128 bits</td>
<td>64 bits</td>
</tr>
<tr>
<td>Rainbow Ia - Cyclic</td>
<td>[27]</td>
<td>Multivariate</td>
<td>58144</td>
<td>92960</td>
<td>64</td>
<td>143 bits</td>
<td>106 bits</td>
</tr>
<tr>
<td>Dilithium IV [30], [31]</td>
<td>Module Learning with Errors</td>
<td>1760</td>
<td>3856</td>
<td>3366</td>
<td>174 bits</td>
<td>158 bits</td>
<td>158 bits</td>
</tr>
<tr>
<td>Falcon 1024</td>
<td>[34]</td>
<td>NTRU</td>
<td>1793</td>
<td>2305</td>
<td>1330</td>
<td>263 bits</td>
<td>230 bits</td>
</tr>
</tbody>
</table>

Finally, Picnic [38] is a NIST candidate signature scheme that does not rely on structure hardness (lattice, quadratic equations, Merkle trees). Picnic depends on zero-knowledge proofs, and symmetric key primitives like hash functions and block ciphers.

### TABLE I: Conventional and Post-Quantum Signature Algorithms and Parameter Sets used in this study.
IV. POST-QUANTUM AUTHENTICATION IN TLS 1.3

This section presents details regarding the implementation and integration of PQ authentication into TLS 1.3 with its implications. To integrate PQ signatures into X.509 and TLS we utilized the OQS OpenSSL [74] library, which is a fork of OpenSSL that introduces post-quantum algorithms from the liboqs library [73, 90]. The OQS OpenSSL version we used was based on OpenSSL version 1.1.1c with ASM optimizations for SHA256, SHA512, RSA and ECDSA256. No ASM optimizations were included for ECDSA384 which we used as reference/comparison in our experiments (see Section V).

Migrating to PQ authentication will require changes in X.509 and SSL libraries; in our case OpenSSL. Fig. 1 shows the format of a PQ X.509 certificate and the fields where PQ algorithm support will need to be added. The new certificate will carry the subject’s PQ public key and the specific PQ signature algorithm used to create the signature. Then the certificate will be signed by the issuer. The PQ signature will be placed in the Signature field. The addition of the PQ public key and PQ signature to the X.509 certificate will increase the size of the certificate and the size of the related certificate chains (Table I). In TLS 1.3, the maximum default size of a X.509 certificate or certificate chain is $2^{34}-1$ bytes and the signature size limit is $2^{16}-1$ bytes. To include new algorithm support, the OQS OpenSSL library defines new X.509 algorithm identifiers [25, 74].

Next, we focus on the TLS 1.3 handshake pieces affected by PQ algorithms. Fig. 2 shows an overview of the messages exchanged by a client who attempts to set up a TLS 1.3 session with a server utilizing quantum-resistant authentication. We assume a classic web scenario where the client is not authenticated. The ClientHello message will negotiate the desired PQ signature algorithm using the signature_algorithms or signature_algorithms_cert extensions. These are essentially lists of algorithm identifiers. The next adjustment for PQ authentication is the PQ X.509 certificate/chain transmission by the server with the ServerCertificate message that will now include PQ certificates. In addition, the server will sign the transcripts of the handshake and transmit a PQ CertificateVerify message that contains a PQ signature. Currently, the maximum default size of this message is 102.4 KB. Additionally, when a certificate chain exceeds 16 KB, TLS utilizes Record Fragmentation [78]. Finally, the client will use the PQ signature algorithm to perform verification operations to the received signatures before sending its own Finished message to end the PQ TLS 1.3 handshake.

V. PERFORMANCE EVALUATION

In this section we present the performance of each PQ signature scheme after its integration into OQS OpenSSL. In our experiments, we assume a classic web scenario where only the server is authenticated using X.509 certificates. All TLS 1.3 handshakes are full 1-RTT mode without PSK resumption. In order to compare PQ authentication in TLS against current algorithms, we utilize the RSA and ECDSA as baselines. For RSA authentication, we use the 3072-bit version that provides 128 bits of security, (Table I). Regarding ECDSA, we utilize the secp384r1 curve which offers 192 bits of security. We note that ECDSA with secp256r1 or Ed25519 would offer 128-bit security level equivalent to RSA3072 with better performance, but we chose higher security level for ECC signatures because it is believed that these primitives will be broken by a QC before their equivalent security level big number (RSA) signatures. We could have tested against RSA2048 or ECDSA256 instead, but our experiments showed that slightly faster and smaller signatures will negligibly speed up handshakes that take an average ~110ms [63]. RSA and ECDSA/EdDSA would offer 0 bits of security in a post-quantum setting. Our objective is to document strictly PQ authentication in TLS. Thus, we utilize a classic X25519 elliptic curve Diffie-Hellman (ECDH) key exchange with Curve25519.

The experiments involved one local machine and several cloud-based instances. The local host was equipped with an Intel i5-8350U processor and 16 GBs of RAM, while the experiments were implemented on a virtual machine that utilized four cores running at 1.7 GHz each and 8GB of RAM. Our cloud machines were Google Cloud Platform (GCP) n1-standard-2 [88] instances running on an Intel Xeon processor with two cores at 2 GHz each and 8 GBs of RAM. All the participating machines were running Ubuntu 18.04 in an x86_64 architecture. TCP Segmentation Offload (TSO) and Large Receive Offload (LRO) were enabled on the hosts’ virtual network cards.
Readers should note that in our evaluation we do not consider the impact of PQ signatures employed in CRL and OCSP revocation checks which would take place over a different connection. In addition, potential use of OCSP staples would introduce another signature generation and verification. Staples can be cached which would not affect the handshake until the staple expired. That falls outside the scope of our work. Similarly, we do not consider SCT checking which requires more signature verifications and TLS data. PQ revocation and SCTs could slow down the handshake, as they do already, but there are techniques to alleviate their impact which we discuss in Section VII-B.

Finally, we use the following metrics to evaluate the TLS performance from the server and client’s perspective:

- **TLS Handshake Time**: The time between the ClientHello message and the end of the handshake. This metric models the latency a client experiences before establishing a secure encryption tunnel excluding the TCP handshake.

- **Certificate Transfer Time**: The time between the ServerHello message and the Client Change Cipher Spec message that includes the transfer of certificate chains from the server to the client.

- **TLS Transactions per Second**: This metric expresses the rate of successful HTTPS connections the server is able to serve.

### TABLE II: Duration of Sign and Verify Operations: Mean and Standard Deviation

<table>
<thead>
<tr>
<th>Signature Algorithm</th>
<th>Local Machine (ms)</th>
<th>Cloud Instance (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sign</td>
<td>Verify</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>St. Dev.</td>
</tr>
<tr>
<td>RSA 3072</td>
<td>3.19</td>
<td>0.023</td>
</tr>
<tr>
<td>ECDSA 384</td>
<td>1.32</td>
<td>0.012</td>
</tr>
<tr>
<td>Dilithium II</td>
<td>0.82</td>
<td>0.021</td>
</tr>
<tr>
<td>Falcon 512</td>
<td>5.22</td>
<td>0.054</td>
</tr>
<tr>
<td>MQDSS 48</td>
<td>10.30</td>
<td>0.147</td>
</tr>
<tr>
<td>Picnic LIFS</td>
<td>4.09</td>
<td>0.050</td>
</tr>
<tr>
<td>SPHINCS+ SHA256-128f-simple</td>
<td>93.37</td>
<td>0.654</td>
</tr>
<tr>
<td>Rainbow Ia</td>
<td>0.34</td>
<td>0.015</td>
</tr>
<tr>
<td>Dilithium IV</td>
<td>1.25</td>
<td>0.021</td>
</tr>
<tr>
<td>Falcon 1024</td>
<td>11.37</td>
<td>0.102</td>
</tr>
</tbody>
</table>

### TABLE III: Sizes of certificate chains (excluding the root), and CertificateVerify extension

<table>
<thead>
<tr>
<th>Signature Algorithm</th>
<th>Cert Chain Size (KB)</th>
<th>CertificateVerify Size (KB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA 3072</td>
<td>1.63</td>
<td>2.44</td>
</tr>
<tr>
<td>ECDSA 384</td>
<td>1.34</td>
<td>2.15</td>
</tr>
<tr>
<td>Dilithium II</td>
<td>6.90</td>
<td>10.42</td>
</tr>
<tr>
<td>Falcon 512</td>
<td>3.54</td>
<td>5.37</td>
</tr>
<tr>
<td>MQDSS 48</td>
<td>42.24</td>
<td>63.42</td>
</tr>
<tr>
<td>Picnic LIFS</td>
<td>66.20</td>
<td>99.57</td>
</tr>
<tr>
<td>SPHINCS+ SHA256-128f-simple</td>
<td>34.46</td>
<td>51.74</td>
</tr>
<tr>
<td>Rainbow Ia</td>
<td>116.86</td>
<td>175.35</td>
</tr>
<tr>
<td>Dilithium IV</td>
<td>10.70</td>
<td>16.11</td>
</tr>
<tr>
<td>Falcon 1024</td>
<td>6.56</td>
<td>9.89</td>
</tr>
</tbody>
</table>

### A. Speed of Cryptographic Operations

First, we evaluate the performance of cryptographic operations for the signature algorithms used in this study. Sign and verify operations are significant differentiation points for the PQ algorithms since they are performed by the server and the client respectively in the TLS handshake. The measurements for the classic algorithms (RSA, ECDSA) were taken using OpenSSL speed command, while for the PQ schemes we used the testing scripts from liboqs [73], [90]. Table II shows the average sign and verification times of the various schemes run on the machines of this study. Since the duration of these operations directly affects the experiment’s TLS handshake time, we measure the algorithms’ standalone performance in milliseconds for each type of machine used instead of CPU clock cycles which is the common practice. We observe that Dilithium, and Rainbow offer competitive performance for 128 and 192-bit security comparable to classical algorithms (RSA3072 and ECDSA384). Falcon offers great verification times, and slower signing. We can see that only Falcon is slower in the cloud machine. We attribute this behavior to Falcon’s floating-point operations which do not exist in other PQ signature schemes. MQDSS and Picnic are slower and SPHINCS+ is significantly slower, especially for sign operations. Our signing and verification measured performance is generally confirmed by the SUPERCOP benchmarks [9] and [83].

NIST Round 2 submissions included optimized versions of the algorithms that take advantage of parallelization and hardware acceleration (e.g., AVX2). These optimizations speed up performance of the algorithms significantly. The implementations we used in our experiments (Table II) were not optimized. Using the optimized versions would make signing and verification greatly faster. For example, AVX2 optimized Falcon signs ~20 times faster and AVX2 optimized Dilithium IV signs faster than Dilithium II without AVX2. We expect that future hardware acceleration of the algorithms themselves would improve their performance further and eventually the TLS cert chain and CertificateVerify size will be the bottleneck of PQ authentication. The results of our work are still relevant for use-cases where acceleration is not available (potentially on constrained devices), but they could underestimate the performance we can get from each scheme which we discuss in the relevant sections in our analysis below.
B. PQ TLS Overhead Analysis - NIST Security Level 1

At the first phase of our experiments we studied the performance of PQ schemes in NIST’s Level 1. The experimental setup consisted of our local machine operating as a client in N. Carolina and a remote machine from GCP’s N. Virginia region acting as the server. Two chain lengths were considered in our experiments as shown in Fig. 5. The first involves a single ICA and the second involves two ICAs since these account for 77% of the SSL cert chains on the Internet today [86].

The same PQ signature scheme was used for all certificates across each chain. Table III summarizes the exact Distinguished Encoding Rules (DER) encoded certificate chain sizes excluding the root CA cert which is not sent to the client, and the TLS CertificateVerify extension size.

Current WebPKI certificates include extensions (SCT, OCSP, CRL, SAN) that we did not use in our experiments. These fields sometimes lead to conventional certificates that exceed 1.5KB. Our tests included less extensions resulting to slightly smaller sizes than common web certificates. Based on the data we collected, we do not expect that one or two hundreds of bytes of certificate extensions would effectively change our results unless they push the certificate chain over the TCP congestion window (cwnd). PQ CertificateVerify and TLS handshake error. After examining the trace, we observed that the client was issuing an SSL Alert because of the certificate public key size, however, the server transmitted the chain before closing the connection. Since this happened consistently for all handshake attempts, Fig. 3 shows the duration of this exchange (the * denotes partial handshake without a ClientFinish message).

Fig. 4 shows the empirical cumulative density function of the single ICA measurements. The long tails observed are attributed to the network transmission path. It is evident from the results that long certificates and slow operations significantly reduce the authentication performance in TLS for algorithms like MQDSS, Picnic, and SPHINCSk. Fig. 6(a) shows the cumulative data transfer between the ServerHello and the Client Change Cipher Spec message for 200 consecutive handshakes providing a insight on the network traffic generated when implementing PQ certs in TLS. Fig. 6(b) shows the average time elapsed during this period. As expected, algorithms with bigger public keys and signatures lead to longer time to transfer the certificates and TLS CertificateVerify extension size.

SCT and OCSP signatures would add significant data to the handshake. We further discuss this point in Section VII-B.

In our performance experiments, we first measure the average TLS handshake time over 1000 handshake attempts. Fig. 3 shows these results for the PQ algorithms in NIST Level 1 for the two certificate chain lengths considered. During our experiments, all TLS records were received successfully and no errors were observed due to the excessive size of the certificates for six out of the seven algorithms examined. More specifically, the TLS client application was not able to process the size of a Rainbow certificate yielding an excessive message size error. After examining the trace, we observed that the server was issuing an SSL Alert because of the certificate public key size, however, the server transmitted the chain before closing the connection. Since this happened consistently for all handshake attempts, Fig. 3 shows the duration of this exchange (the * denotes partial handshake without a ClientFinish message).

Fig. 4 shows the empirical cumulative density function of the single ICA measurements. The long tails observed are attributed to the network transmission path. It is evident from the results that long certificates and slow operations significantly reduce the authentication performance in TLS for algorithms like MQDSS, Picnic, and SPHINCSk. Fig. 6(a) shows the cumulative data transfer between the ServerHello and the Client Change Cipher Spec message for 200 consecutive handshakes providing a insight on the network traffic generated when implementing PQ certs in TLS. Fig. 6(b) shows the average time elapsed during this period. As expected, algorithms with bigger public keys and signatures lead to longer time to transfer the certificates and TLS CertificateVerify extension size. For small signatures (i.e., Dilithium II, Falcon 512) that equates to less than 5ms extra whereas for bigger ones (i.e., MQDSS, Picnic, SPHINCSk) it
becomes ∼35-55ms extra.

The client and server were in close proximity in this experiment. The transfer times of lengthy signatures and chains could exceed 0.5s for longer client-server distances. That is because algorithms with excessively large certificate sizes and signature messages are penalized by the TCP congestion window. Certificate chains and CertificateVerify payloads of high size (i.e., MQDSS, Picnic, and SPHINCS+) from Table III incur round-trip(s) which significantly slows down the handshake due to the TCP congestion window (cwnd). Falcon 512 and Dilithium II, on the other hand, do not incur extra round-trips and the additional cert transfer time would remain ∼5-10ms even over longer distances.

In Fig. 3 algorithms that result to smaller certificate chains and CertificateVerify payloads (i.e., Dilithium, Falcon from Table III) show competitive performance against conventional signature algorithms. Falcon 512 has smaller signatures and public key, but we see it performing worse than Dilithium II because its signing operation at the server takes longer (∼6.5ms without AVX2 optimizations). Signing duration of 6.5ms and chain/TLS signature transfer of ∼3ms are more noticeable in a localized scenario (e.g., a server in N. Virginia) with a baseline RSA3072 handshake time of ∼15ms. Falcon 512 would look better for the average Internet web connection of ∼110ms [63] as we show in Section V-E.

C. PQ TLS Overhead Analysis - NIST Security Levels 3, 5

Next, we extend our experiments to measure the performance of algorithms that belong to higher security levels in NIST’s level scale. For the majority of the examined PQ algorithms, the parameter sets that provide higher quantum-resistant authentication result to significantly larger public keys and signatures [3, 5, 21, 27, 31, 34, 38] and worse performance. Due to these sizes, the certificate chain and CertificateVerify message were shown in Section V-B to exceed the TCP initcwnd which causes significant slowdowns because of the extra round-trip(s).

Thus, for Levels 3 and 5, we limit our attention to the two PQ algorithms that presented the most promising results at Level 1, namely Dilithium II and Falcon 512. We examine Dilithium IV at NIST security Level 3, and Falcon 1024 at Level 5. Currently, Dilithium does not offer a parameter set at Level 5, while the Falcon’s Level 3 parameter set (Falcon 768) was removed in Round 2 [34].

Again, we measure the average TLS handshake time over 1000 handshakes for our client in N. Carolina and a server in N. Virginia. Fig. 7(a) shows these results for the examined algorithms for the two certificate chain lengths, while Fig. 7(b) shows the empirical cumulative density function of the single ICA measurements. We observe that by increasing the security level of the PQ schemes the absolute added latency to the TLS handshake is ∼10-15ms over RSA3072. Fig. 8(a) shows the cumulative data transfer between the ServerHello and the Client Change Cipher Spec message for 200 consecutive handshakes. Fig. 8(b) shows the average time elapsed during this period which is an additional ∼5ms for Falcon 1024. The Dilithium IV cert transfer is slower because the data size triggers an extra round-trip (∼11ms). Falcon 1024 does not add an extra round-trip which means that the additional cert transfer time would remain ∼5-10ms even over longer distances, but it adds ∼14ms of signing. An additional 14ms is more noticeable in a localized scenario (N. Carolina to N. Virginia) with a baseline RSA3072 handshake time of ∼15ms.

D. Combining PQ Signature Schemes

The diverse nature of the cryptographic primitives used for each PQ signature algorithm leads to solutions with different characteristics and behaviour, namely different signature/key sizes and computational complexity (Tables I, II). Therefore, by using different PQ signature schemes in the same certificate chain, we can leverage each algorithm’s specific strengths and effectively reduce the overall TLS handshake time. To the best of our knowledge, this is the first time where using different algorithms along the certificate chain has been proposed to improve total handshake performance. To illustrate this intuition, we examine a proof-of-concept combination of the Falcon 1024, and Dilithium IV algorithms across a chain with one ICA.

The root CA and ICA certs include a Falcon 1024 public key and signature from the root CA, while the server cert utilizes a Dilithium IV public key and a Falcon 1024 signature from the ICA. By doing so, we reduce the duration of the server’s signing operation by approximately 30 times. In addition, the overall certificate chain size is equal to 6.52 KB, which is smaller than both the pure Falcon 1024 chain (6.56 KB), and the pure Dilithium IV chain (10.70 KB) (Table III). Table IV shows a comparison yielded following the experimental procedures of Section V-C (client-server round-trip ∼11ms). The combined case reduces (on average) the overall handshake by 25.16% compared to the single Dilithium IV, and by 33.27% when put against the single Falcon 1024 case. The combination can work with other signature schemes too as discussed in Section VII-B. It is especially suitable for the cases where (a) the client and server are in close proximity, and (b) key and signature sizes are penalized by the TCP congestion window.
proximity, (b) sufficiently fast implementations of the signature scheme are not available at the signer or verifier, or (c) when leaf certificates are renewed relatively often and lower security levels with better performing parameters are acceptable.

The drawback is that the verifier is required to support both signature algorithms. The performance improvement of this method will be marginal for long client-server distances. The client and server Section V-D and V-E further examine this proposed alternative of combining schemes.

### E. Global Scale Performance Analysis

Next, we extend our evaluation by adding more remote servers across different continents. Fig. 9(a) shows the experimental setup, the average number of hops, and the measured average round-trip latency between our local client and each remote server. Again, the handshake duration was averaged over 1000 handshakes for each PQ signature algorithm and parameter set. A certificate chain with a single ICA was utilized. Fig. 9(b) shows the mean and standard deviation of the TLS handshake time when algorithms with NIST’s security Level 1 were tested, while Fig. 9(c) does the same for Level 3 and 5. The performance was similar across all regions, and no handshake failures were reported due to middlebox misbehaviour. Consistently, Dilithium and Falcon show the most promise for replacing RSA3072 with minimum additional delay.

Using the same setup, in a second experiment, our client uniformly performed 3000 handshakes with the remote servers in the course of a whole day for each examined algorithm. The goal was to account for Internet’s unpredictability by measuring TLS handshake latency in diverse topologies across longer time-frames. Table V summarizes the additional latency introduced to the TLS handshake by the PQ signature algorithms when compared against RSA3072 at the 50th and 95th percentile. The observed latency overhead of Level 1 Dilithium II was notably small making it the most pertinent for integration into TLS. Level 1 Falcon 512 also showed small overhead over RSA3072. Regarding higher security levels, Dilithium IV nearly doubled the handshake duration because the data size triggered an extra round-trip, while Falcon performed better. Although the % increase seems high for Dilithium IV, the absolute time increase is ~145ms at the 50th percentile and ~229ms at the 95th percentile. The Dilithium+Falcon combination (Section V-D) reduced the handshake latency by ~8-50% in comparison to single PQ algorithm certificate chains.

Our experiments so far have measured the time between the ClientHello message and the end of the handshake. Extending the period under examination to include the TCP handshake (start at the TCP SYN message) increases the mean TLS handshake with X25519 key exchange and RSA3072 certificates duration to 269.5ms. The respective median of the handshake measured in Firefox client-server connections is 111ms considering 9.91 billion data points from their nightly versions 69, 70, and 71 (excluding 32-bit platforms which could deteriorate the performance unrealistically) [63].

---

**Table IV: Average Handshake Time and σ of the Dilithium and Falcon combination (single ICA).**

<table>
<thead>
<tr>
<th>Signature Scheme</th>
<th>TLS Handshake (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>RSA 3072</td>
<td>15.13</td>
</tr>
<tr>
<td>Dilithium IV</td>
<td>24.20</td>
</tr>
<tr>
<td>Falcon 1024</td>
<td>27.14</td>
</tr>
<tr>
<td>Fal. 1024 &amp; Dil. IV</td>
<td>18.11</td>
</tr>
</tbody>
</table>

**Table V: TLS Handshake Times and Additional Latency over RSA at the 50th and 95th Percentile.**

<table>
<thead>
<tr>
<th>Signature Algorithm</th>
<th>Handshake (ms)</th>
<th>Latency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50th</td>
<td>95th</td>
</tr>
<tr>
<td>RSA3072</td>
<td>131.54</td>
<td>227.26</td>
</tr>
<tr>
<td>Dilithium II</td>
<td>140.20</td>
<td>232.51</td>
</tr>
<tr>
<td>Falcon 512</td>
<td>142.22</td>
<td>235.46</td>
</tr>
<tr>
<td>MQDSS 48</td>
<td>598.61</td>
<td>726.20</td>
</tr>
<tr>
<td>Picnic LFS</td>
<td>634.90</td>
<td>985.88</td>
</tr>
<tr>
<td>SPHINCS+ SHA256-128f-simple</td>
<td>553.15</td>
<td>904.98</td>
</tr>
<tr>
<td>Dilithium IV</td>
<td>276.55</td>
<td>449.88</td>
</tr>
<tr>
<td>Falcon 1024</td>
<td>152.96</td>
<td>240.74</td>
</tr>
<tr>
<td>Fal. 1024 &amp; Dil. IV</td>
<td>140.74</td>
<td>228.42</td>
</tr>
</tbody>
</table>
extrapolation we estimate that the average PQ TLS handshake measured from the TCP SYN to Ready for HTTP will amount on average to \(\sim 122\text{-}135\text{ms}\) for the best algorithm cases of each security level, namely Dilithium II (level 1), and Falcon 1024 (level 5).

The PQ signing and verification times were relatively small (<15ms) in our analysis so far for all tested algorithms except SPHINCS\(^\dagger\). Such times have relatively little impact on an average handshake time of \(\sim 110\text{ms}\) or \(269.5\text{ms}\) in our tests. Arguably using optimized or hardware accelerated versions of these algorithms would speed up the handshake, but not by a lot because the certificate and signature data will still incur a slowdown.

### F. Server-Side Performance of TLS PQ Authentication

Up to this point, we focused on measuring the impact of PQ authentication in TLS from the client’s perspective. Next, we measure the performance of a server that utilizes PQ certificates in TLS 1.3 to service simultaneous secure connections with multiple clients. To do so, an Nginx web server \([66]\) was set up on one of the aforementioned cloud instances, and was configured to utilize the OQS OpenSSL library. Thus, our server was able to sign handshakes and send PQ certificates using the signature schemes under study. We used the Siege tool \([35]\), configured with the OQS OpenSSL library, to simulate multiple clients and set up PQ authenticated TLS connections. Siege clients were running in four remote cloud instances. For this experiment, all client instances were placed in relatively close proximity to each other. The exact topology details are shown in Fig. 10.

The clients that were uniformly distributed among the four locations, attempted simultaneous TLS connections with the server for 60 seconds. During that time we measured the request rate that the server was able to handle, i.e. the number of successful transactions per second. The requested web page was only 0.6KB. In addition, we captured the server’s overall availability by measuring the number of TLS handshake failures during the load testing period. Fig. 11 shows these results for promising signature algorithms of NIST’s Level 1 as the total number of clients increases from 20 to 1000. Both cases of one and two ICAs are examined. Evidently, while still at low load RSA3072 outperforms Dilithium II and Falcon 512 as expected from Fig. 5. However, as the server’s saturation point is reached we observe that Dilithium II allows the server to handle \(\sim 25\%\) more connections per second compared to RSA3072 because Dilithium signing is over five times faster compared to RSA3072 at the server (Table II). Regarding Falcon 512, the server’s saturation point is reached early at a much lower transaction rate than RSA and Dilithium due to slow signing at the server (Table II).

Fig. 3 showed that the Dilithium bigger cert chain and CertificateVerify size do not affect the server performance other than slightly slowing down the handshake. To further reinforce that point, the server’s transaction rate is similar for both cert chain lengths (Fig. 11(a) vs. Fig. 11(b)) for the same algorithms which shows that just a few additional KB of certificate payload do not affect the server performance.

To further demonstrate the higher impact of signing over cert chain size on a busy server we tested Picnic LIFS which presents much bigger certificate chain and CertificateVerify message (Table III) while maintaining a better than Falcon 512 sign operation (Table II). Fig. 11 shows that the server signing with Picnic LIFS outperforms signing with Falcon 512 by \(\times 1.7\) higher transaction rate. That can be attributed to the sign operation performance. Moreover, and in accordance with our earlier experiments, Picnic still leads to a much lower maximum transaction rate compared to RSA3072 and Dilithium II due to its excessively big cert chains that slow down the handshake significantly (Fig. 3). Another interesting observation is that before reaching the server saturation point, Picnic’s transaction rate is half than RSA3072. The reason is that up to this point the server has not gotten overloaded and the transaction slowdown solely comes from the TLS performance slowdown seen in Fig. 3.

The same behaviour is observed for parameter sets that yield higher NIST security levels as seen in Fig. 12. The server’s saturation point is reached significantly early with Falcon 1024 that also presents the highest handshake failure rate among alternatives because of its signing performance.

---

**Fig. 10:** Round-trip latency and number of hops between clients and server

**Fig. 11:** Server Performance for NIST’s PQ Security Level 1 Signature Algorithms: (a) One ICA, (b) Two ICAs

**Fig. 12:** Server Performance for NIST’s Higher PQ Security Level Signature Algorithms: (a) One ICA, (b) Two ICAs
Also, under high load, the use of Dilithium IV results to higher transaction rates in comparison to RSA3072 and Falcon 1024 because of its better signing performance (Table II). Moreover, we examined the proposed PQ combined certificate chain scenario (Dilithium and Falcon in Section V-D). Results demonstrate that by using PQ combined certificate chains the server is able to handle more TLS connection requests than all alternatives including the conventional RSA3072. More specifically, the Dilithium+Falcon combination slightly increases the average post-saturation transaction rate by ∼10% and ∼4% compared to RSA3072 and pure Dilithium IV respectively. The improvement over RSA3072 is because Dilithium IV operations are still faster. However, the difference is smaller than in Fig. [1] because the Dilithium IV chain slows down the handshake more (Fig. [7]). The server’s performance increase with Dilithium+Falcon over Falcon 1024 is due to the computationally lighter signaling of Dilithium IV (Table II).

In summary, we saw that signing at the server is an important factor for the server’s performance. Cert chain and signature size affect the server transaction rate by increasing the TLS handshake time but the effect is smaller than the impact of a heavy signing operation. We expect that the server performance penalty introduced by expensive sign operations will be alleviated by optimized implementations and hardware acceleration in the future.

Finally, it is worth noting that in our tests we observed that the client closest to the server saw a higher transaction rate due to the lower propagation delay. Thus, servers closely located to clients will see better performance than the average we are showing.

VI. RELATED WORK

A large body of research on QC and PQ cryptography [11], [44], [50], [55] led to the NIST initiative to come up with novel PQ algorithms for use in a post-quantum world. Recently, more studies are exploring NIST’s Round 2 PQ candidate schemes focusing mainly on their computational performance, while some even evaluate their energy consumption [82]. The authors in [6], [89] conduct a hardware implementation and comparison of NIST Round 2 signature and KEM candidate algorithms. Their focus is on comparing specific algorithm operations and their impact on hardware design parameters, without taking into account performance as a part of a system or a protocol. In [48], Kannwischer et al. benchmark Round 2 algorithms on ARM Cortex-M4 and evaluate the more suitable ones for embedded devices without studying additional impact on specific use-cases. Similarly, the authors in [85] enhance the DTLS protocol with the PQ key exchange algorithm NTRU which showcases its feasibility for securing current systems.

In addition, research teams from the industry are investigating the performance of PQ candidate algorithms in current Internet protocols. They have mostly been focusing on studying key exchange. In [15], Bos et al. proposed the use RLWE for key exchange in TLS 1.2 and study its impact that slows down performance by a factor of ×1.4 for the client and ×1.2 for the server. A year later, in 2016, Google tested a hybrid KEM (NewHope combined with elliptic-curve schemes) and ECDSA certs with TLS 1.2 in Google Chrome [54]. Following the conclusion of NIST’s Round 1 evaluation, Google and Cloudflare resumed the experiments on KEM integration into TLS. The new study [56] defined three KEM families based on the primitive and the key size and tested them with dummy extensions of proportional sizes in the TLS ClientHello message. The goal was to emulate the overhead that these key sizes introduce and explore issues or failures related to such larger handshakes. Recently, Google and Cloudflare have concluded work on a new experiments (CECPQ2 [55]) that integrated a hybrid KEM based on the HRSS NIST PQ candidate and X25519 into TLS 1.3 for further performance measurements on Chrome. Additionally, they tested the SIKE NIST PQ candidate with X25519 in TLS 1.3 as part of their CECPQ2b experiment [52]. The results of this work are included in [53]. Moreover, Campagna recently discussed hybrid key exchange and double tunnelling for TLS and presented the slowdown of TLS 1.2 introduced by using hybrid key exchange with ECDHE and SIKE or BIKE NIST PQ candidates [18]. Campagna et al. showed that SIKE’s impact on the handshake byte count is small, whereas BIKE’s impact is significant.

More closely related to our work that focuses on signature performance, Cronin et al. [25] modeled and benchmarked the performance of forward-secure signatures against conventional RSA, DSA and ECDSA. Focusing more on PQ signatures, Kampanakis et al. discussed the impact of PQ signature schemes on protocols that utilize X.509 certificates in [36]. They considered the case of hybrid certificates that include hash-based signatures, and conducted performance experiments on TLS 1.2 and IKEv2. Also, their work discussed the use of PQ certificates by emulating their increased size through enlarging certificate chains with additional certificates. On the same topic, earlier work by Bindel et al. emulated large hybrid PQ certificates and studied their impact on TLS libraries and browsers [13]. The authors in [17] prototyped stateful XMSS signatures in TLS and S/MIME and pointed out the challenges with state management in live TLS connections. In [25], the authors discuss the challenges of implementing NIST’s PQ key exchange and authentication algorithms in TLS, with their focus being mainly on hybrid (combination of PQ and conventional) schemes. In addition, they perform proof-of-concept experiments on PQ TLS that involve single certificate exchanges, without accounting for real-world network conditions. Finally, Paquin et al. [69] perform experiments with some hybrid PQ KEM and Signature schemes in TLS and show that lossy conditions have more negative effects on data-heavy TLS handshakes. The loss rates they use are not very common on the Internet today, but they could be experienced in constrained or congested environments, or remote cellular networks.

VII. DISCUSSION ABOUT ENCRYPTED TUNNELS

A. Implications of our Findings

PQ signatures will have an impact on authenticated tunnels like (D)TLS and IKEv2/IPSec. These protocols provide fragmentation mechanisms to allow for lengthy signatures, but as we showed above, larger chains and potentially slower algorithms will impact the tunnel establishment. Applications with lower connection rates and tunnels that stay up longer will be less impacted than applications that establish short and fast connections. Thus, per connection overhead is important to be
considered for the migration to PQ authentication algorithms. In addition, the more signatures used in a handshake, the more impact PQ algorithms would have on the protocol. For example, SCTs in leaf certificates will add two or three signatures to the handshake and OCSP stapling would add one more.

IPSec VPN tunnels usually stay on for long periods of time. Experimental RFC4478 [67] defines how a tunnel can be re-authenticated, but in practice, IKE_AUTH messages are not exchanged unless the tunnel is torn down and re-established. Thus, for the majority of IKEv2/IPSec VPN applications, a connection establishment that would take up to a few seconds would not have material impact on the tunnel. Similarly, WebVPN applications establish a TLS control connection and subsequently data DTLS connections over a duration of a few seconds. Onward, these tunnels stay up for long sessions that usually last hours or more. A connection establishment slowdown of up to a few seconds will not impact these connections. The caveat with such use-cases is the impact on the head-end that terminates these connections. It is not unusual for a VPN concentrator to terminate thousands of connections. A possible attempt of VPN clients to establish a connection to the hub simultaneously could lead to overload if the signing or verification operations are heavy. Most PQ signature candidates we studied (except SPHINCS\textsuperscript{+}) are not likely to be a significant performance concern for VPN tunnels that take advantage of algorithm optimizations.

Web connections, on the other hand, are usually short-lived. At the time of this writing, clients perform 70 requests per page to fetch 1-2KB resources per request on average [42]. If some of these requests are performed over new TLS connections, a heavier PQ authenticated TLS handshake would have a significant impact on HTTP performance overall. A few KB of extra authentication data per TLS connection has low amortization over 1-2KB of actual web content. On the other hand, the ever-increasing adoption of HTTP/2 [7] (55% at the time of this writing [22]) will improve amortization as HTTP/2 multiplexes data over a single HTTP connection. It is hard to say with certainty how much delay is excessive on the web. Reports like [1] lead us to believe that hundreds of extra milliseconds per handshake are not acceptable. More importantly, extra round-trips mean almost doubling total handshake times which we consider excessive for time-sensitive web applications.

In summary of our testing, we found that Dilithium II, Falcon 512 and 1024 could be deployed in X.509 certs without detrimentally impacting time-sensitive TLS applications. Dilithium is preferable over Falcon, at least when floating point hardware is not available, because its superior signing performance allows for higher connection rates. Floating point hardware improves Falcon’s signing performance by \(\sim 20\) times [72] which would make Falcon a better candidate than Dilithium in terms of performance in live protocols. Falcon would also have more impact on energy constrained devices [83]. Dilithium IV nearly doubled the handshake duration over RSA3072 because the data size triggers an extra round-trip. Dilithium IV could still be used if applications were amenable to approximately double the TLS handshake time or if some of the mechanisms in Section VII-B (e.g., [92], [79] § 5.1.3) were widely deployed. We also showed that other NIST PQ signature candidates would not be good candidates for such applications.

Dilithium II and Falcon 512’s claimed classical security levels are \(\sim 100\) and \(\sim 114\) bits respectively. Although these asymptotic bounds are probably higher in practice because of memory costs, these parameter sets could be considered unacceptable to use. The variants that offer more than 128 bits of classical security are Dilithium III, IV and Falcon 1024. We already discussed that Falcon 1024 had acceptable handshake time with the shortcomings of higher energy cost and slow signing when floating point hardware is not present and that Dilithium IV was introducing extra round-trips. A rerun of our experiments in Section V-E showed that the use of Dilithium III certificates added 2.7KB to the handshake (compared to Dilithium II) with fast signing/verification times. That led Dilithium III to perform \(\sim 7\)% slower than RSA3072 which is very similar to Dilithium II. Based on the above, if Dilithium II and Falcon 512 security levels are considered low, Dilithium III and Falcon 1024 are the best options. Falcon 1024 would be preferable when floating point hardware is available at the signer and energy is not a concern.

None of our tests included OCSP revocation checks and SCTs returned from the server. OCSP and SCTs could account for three or more PQ signatures which could push the server TLS data over the TCP initcwd and introduce round-trips for all Dilithium variants. Falcon data could fit in the initcwd at the cost of one slower OCSP signature assuming floating point hardware is not available. Falcon would be preferable specifically for SCTs because of its relatively fast verification and small signature. Readers should note that on the web, OCSP responders are often not used by browsers [58] and staples can be cached and are not supported by most servers. Thus, we consider OCSP signatures a less pressing challenge for the PQ WebPKI case.

B. Minor Adjustments to Enable PQ Signatures

There are straightforward changes needed for TLS and IKEv2 to support PQ signatures. The TLS extensions signature_algorithms_cert and signature_algorithms will need to use new identifiers to convey the TLS peer the PQ signature algorithms supported [25] § 4.1.1. Also, X.509 will need to use new Public Key and Signature identifiers to convey the new algorithms [25] § 4.1.1, [46] § 1. In the context of IKEv2, signature negotiation was not included in the protocol from the beginning. Recently RFC7427 [49] added support for signature algorithm negotiation. RFC7427 is rarely supported by VPN vendors, so in a PQ authenticated IKEv2 case, vendors will have to support all possible algorithms and not negotiate with the peer, or support RFC7427 and negotiate the PQ signature algorithms with new identifiers.

Crockett et. al. describe in [25] § 4.3.1 the challenges introduced by some algorithms with signatures and cert chains that exceed 2\textsuperscript{14} B and 102.4 KB respectively. Out of the tests we performed, only Rainbow fell in these categories and led to failed handshakes. We do not expect schemes with such big signatures to be used in TLS, thus TLS and protocol implementations will likely not require further updates to accommodate them.
On the other hand, the work in \cite{92} proposes a new TLS extension to inform the server that the client does not need the ICA certificates in the certificate chain. The authors in \cite{99} § 5.1.3] propose a similar method of omitting certificates from a handshake by using a pre-established certificate dictionary. Such mechanisms could save significant amounts of data from handshakes that approach or exceed the TCP initcwnd to prevent extra round-trips similar to Dilithium IV. They would also alleviate the impact of lossy networks which was shown to be high for data-heavy TLS handshakes.\cite{69}. The way to implement the mechanism would be either for the client to keep track of ICAs, or maintain an ICA cache and a server leaf certificate Common Name (CN) cache. Both caches would get updated from a new TLS handshake when there is a cache miss. The server certificate itself would not be cached. It would still be returned from the server with every handshake. The CN cache could save additional data overhead from OCSP and SCT PQ signatures as well. Valid OCSP staples and SCTs could be bound to a CN cache entry. As long as the client had an entry in the cache for the server leaf certificate, he would not need to receive the ICAs, OCSP or SCT information. One challenge would then be that, at the time of this writing, SCTs are most usually included in the leaf certificate. To omit the SCT data, the certificate could no longer include the SCT information which would now be carried in a signed_certificate_timestamp extension\cite{57}. That is a drastic paradigm change but it would allow the client to limit the TLS handshake data if the information is in its cache.

Section \ref{V-D} showed how combining algorithms could improve the performance of the handshake. Using the same rationale at the root CA, the algorithm combination with relatively small signature size could shrink the ICA certificate size and speed up the handshake. Such schemes could be multivariate candidates like LUOV, GeMSS, and Rainbow or, qTesla, or Stateful HBS\cite{43,60} with small tree heights used at the root CA. Certificate chains that include different signature algorithms impose an extra requirement for a client to support multiple PQ signature algorithms.

\subsection*{C. Drastic changes to Enable PQ Authentication}

CPU-intensive algorithms like Falcon signing could add significant load to busy servers. Such cases could benefit from batch signing options proposed in \cite{8}. Batch signing would allow the server to sign batches of handshakes with one signature, but it would require broad client upgrades and could introduce delays while the server is buffering handshakes to batch-sign them.

Some lattice signature schemes like Falcon could offer message recovery as part of the signature\cite{71}. If the message is longer than a plain Falcon signature then message recovery shortens the \texttt{message+sig} size. This mechanism could be integrated in X.509 to make the PQ certificate smaller. An X.509 certificate could then only include the algorithm identifier and the signature. The message \texttt{(tbsCertificate)} would be recovered from the signature itself. Such methods would require significant changes in X.509 standards and their implementations in tunnelling protocols.

If Dilithium and Falcon are not standardized, tunnelling protocols will be significantly affected by the rest of the PQ signature algorithms. Given that the industry is constantly striving for faster handshakes and better performance\cite{80,92,79} it is unlikely that the impact of the PQ signature algorithms (excluding Dilithium and Falcon) will be acceptable. In that case, more drastic protocol changes may be necessary.

One thought proposed in public fora is to follow the Certificate Transparency paradigm\cite{57} and offer a public repository or service that could be used out-of-band to retrieve certificates of the entities that a client would communicate with. To offer the same functionality, the work in \cite{80} uses special DNS Resource Records to serve public keys used for SNI encryption. Such proposals would eliminate the need for certificates to be transferred in the handshake, but it comes with a requirement of retrieving and verifying the peer certificate out-of-band. Unless caching is enforced, this certificate retrieval step would need to be repeated before any TLS connection establishment which of course is unlikely to improve the overall performance. If caching the peer certificate is possible, it comes with concerns\cite{84} § 7, \cite{46} § 2.1]. We consider such a mechanism a drastic paradigm shift from the way the Internet works today.

On the other hand, the authors in \cite{14} Figure 2, \cite{10} § 2.3, \cite{81} propose an alternative authentication method that uses PQ Key Encapsulation which could be more efficient than PQ signatures. The mechanism would require for the server leaf certificate to include a PQ KEM public key which can be relatively short. But it would also require new protocol message extensions and an extra round-trip in TLS because the signature ciphertext can only be generated after the client has retrieved the server's public key. It would work similarly for IKEv2. Changing the TLS or IKEv2 state machine in order to prevent the extra round-trip would not be trivial as it would affect current security analyses of the protocols\cite{23,24}. Moreover, the public key would still need to be tied to the peer identity using PKI so the PQ cert chain containing PQ signatures would still need to be transferred which minimizes the improvement this mechanism offers.

\section*{VIII. Conclusion and Future Work}

In this work, we integrated PQ signature algorithms into TLS 1.3 and evaluated the TLS handshake latency observed by a client along with the throughput of a PQ authenticated server by considering realistic network conditions. We proved that the signature and certificate chain size impact the total handshake time, especially with relatively fast signing and verification primitives. Our results show that the PQ algorithms with the best performance for time-sensitive applications are Dilithium and Falcon. When floating point hardware is available Falcon seems more suitable for the web. Other applications (e.g. VPN, SSH) that do not require frequent connection establishments could use one of the other PQ signature algorithms as well, assuming they do not overload a server terminating multiple client connections. We showed that although slightly slower signing does not significantly affect the overall duration of a single handshake, it could significantly impact the total throughput of a server. However, as optimizations and hardware acceleration improve signing performance, we expect signature and key size to have the most impact on the handshake. We also proposed combining different PQ signature algorithms in a certificate chain and
we confirmed that it can improve the overall handshake speed and throughput. Finally, we offered improvements to avoid round-trips by leveraging ICA suppression and discussed other challenges and alternatives in real-world protocols.

As future work, we plan to evaluate the performance of PQ authenticated VPNs and UDP-based tunnels like QUIC and DTLS [17]. We also want to test hybrid PQ KEMs as tested by others [2,3,5] along with PQ signatures to evaluate the "total" latency introduced to TLS by PQ algorithms. By combining results in [5] with our work, we could extrapolate that the total slowdown of a PQ handshake would be around 10-25%. Also, given that PQ key exchange could add 1-2KB each direction, this could push the server data to the initcwnd limit and introduce round-trips. Further experiments would be necessary to quantify the total PQ algorithm impact on TLS under realistic conditions that include lossy networks. It would also be of interest to study the impact of PQ SCT and OCSP signatures while using the caching and ICA suppression mechanisms already discussed. Finally, it is worth investigating hybrid certificates' performance and the improvement that Stateful HBS, qTesla, GeMSS, LUOV or Rainbow would achieve when used only at the root CA. As a longer-term goal, in an attempt to reduce the certificate size, we would also like to study the message recovery capabilities offered by schemes like Falcon.

ACKNOWLEDGMENTS

Many thanks to Richard Barnes from Cisco for his valuable feedback. We would also like to acknowledge Luke Valenta from Cloudflare for his comments and experimental results regarding longer TLS records in TLS handshakes. Hanno Böck also was kind to challenge us insightfully on the concerns PQ signatures bring to PKI infrastructures. We are grateful to Thom Wiggers and Prof. Lejla Batina from Radboud University for their thoughtful feedback which improved this paper. Finally, thank you to Douglas Stebila, Christian Paquin and the whole OQS OpenSSL [74] team for providing a library that made our testing possible.

REFERENCES


[25] E. Crockett, C. Paquin, and D. Stebila, “Prototyping post-quantum and
[62] Nginx, “NGINX: High Performance Load Balancer Web Server and...


