# (Yet another) Analysis of MLWE's performance impact on specific TLS Use-cases

Panos Kampanakis[1] and Alex Weibel[1]

[1]Amazon Web Services
{kpanos}@amazon.com

June 15, 2025

***Abstract**: The overall goal of this work is to get closer to a conclusion which confirms or disproves the claim that the size of MLWE will significantly affect most common applications running over TLS. It does so by building on the qualitative results in [6] and measuring actual web performance metrics of real-world web content.*

It is sometimes argued that post-quantum cryptography will have significant performance impact on remotely connected applications. A common example is quantum-resistant cryptographic algorithms used in TLS 1.3 for key exchange and server authentication for the Web. Although the new algorithms have processing performance similar to today's classical algorithms, they introduce kilobytes of additional data in the connection setup which ends up slowing down TLS 1.3 handshakes. TLS 1.3 was designed for security and speed. In its most typical use, it can start encrypting data from the server to the client after one round-trip (RTT). Figure 1 shows a typical, simplified TLS 1.3 handshake from [19]. Using the newly standardized quantum-resistant Module Learning With Errors (MLWE) algorithms for key exchange (ML-KEM) and authentication (ML-DSA) could add ∼1.5KB in the `ClientHello` and over 15KB in the `ServerHello`, `Certificate`, and `CertificateVerify` TLS 1.3 messages compared to their classical ECDH and ECDSA conterparts. Although this data could slow down the TLS handshake, will it materially impact an application's performance?

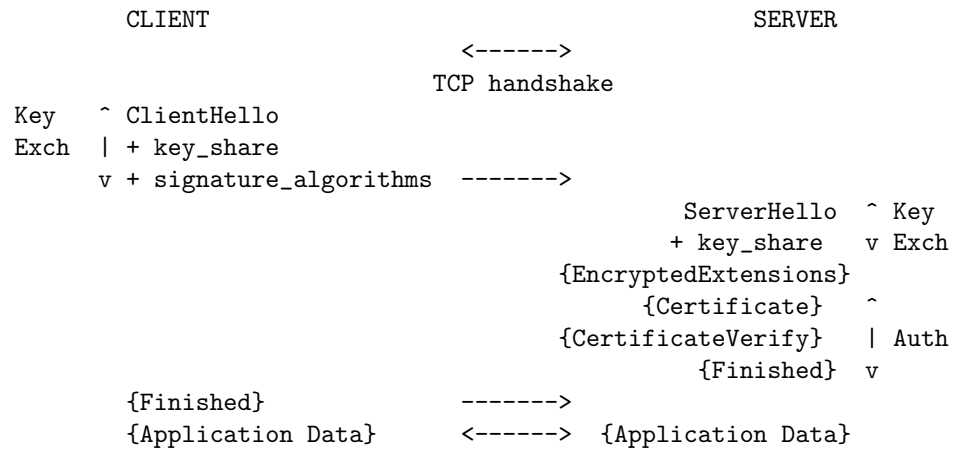



Figure 1: A typical TLS 1.3 handshake [19] with application data flowing after two RTTs.

This work expands on previous studies that investigated the topic. Many previous analyses [12, 13, 11, 22, 17, 21, 18, 1, 3, 14, 4] laid out the issue with additional data especially under more constrained networks. They showed that MLWE algorithms will not affect the transport protocol when the networks is fast and stable, but when the bandwidth is limited or under lossy or constrained conditions, the additional data will affect the 90-th percentile and can add over 20% to the handshake completion time. They showed that it initial TCP congestion window of 10 Maximum Segment Size units could ntroduce a round-trip unless the server has increased it for better performance. They also demonstrated that network middleboxes can sometimes cause failures for larger ML-KEM-768 TLS 1.3 `keyshare`s. To address the quantum-resistant data size issue, various works have proposed to slim it down [9, 20]. Similar concepts have been proposed in the IETF for cutting down the authentication data [7, 2, 5].

Since then, some works have suggested focusing on other metrics that evaluate the performance impact of MLWE key exchange and authentication. NIST National Cybersecurity Center of Excellence's (NCCOE) Post-quantum Migration Project [15] published a draft report [16] which teases out the need for evaluating the time-to-last-byte (TTLB) as a performance metric. This work also showed that QUIC congestion control, amplification protection and packet pacing can introduce more significant slowdowns to a QUIC handshake. [10] proposed the TTLB as a better performance metric. [8] looked into the TTLB and evaluated MLWE's imapct on the TTLB of connections transferring large amounts of data. It showed it to be much lower than the impact on TLS handshakes especially under high or medium speed connections. Intuitively, the reason is that an extra 15KB per handshake does not impact the TTLB of a connection transferring data of $\sim 10\times$ that size. **Full-disclosure**, one co-author of these works is also a co-author of this manuscript.

Recently, Cloudflare [23] shared some data on the web connections they see. They showed that session resumptions are a significant part of their web connections. These will not be affected by MLWE, but for the rest, they noticed that the median of the connection data is small which intuitively means that they will not significantly amortize the cost of the additional MLWE handshake data. [6], on the other hand, explored popular website performance metrics and showed that they transfer large amounts of data. It calculated MLWE's impact on common web performance metrics like TTLB, Largest Contentful Paint (LCP), First

| Website | Standard ML-DSA 4G | | | | | Trimmed ML-DSA 4G | | | | | DC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | TLS +% | TTFB +% | FCP +% | LCP +% | DC +% | TLS +% | TTFB +% | FCP +% | LCP +% | DC +% | [KB] |
| google.com | 7.8 | 1.72 | 1.11 | 1.11 | 0.55 | 5.2 | 1.14 | 0.74 | 0.74 | 0.37 | 445 |
| youtube.com | 7.8 | 1.76 | 1.15 | 1.11 | 1.27 | 5.2 | 1.18 | 0.77 | 0.74 | 0.84 | 62 |
| outlook.office.com | 7.8 | 1.08 | 1.02 | 0.98 | 1.02 | 5.2 | 0.72 | 0.68 | 0.66 | 0.68 | 170 |
| facebook.com | 7.8 | 1.30 | 1.14 | 1.14 | 0.84 | 5.2 | 0.87 | 0.76 | 0.76 | 0.56 | 30 |
| docs.google.com | 7.8 | 0.90 | 0.68 | 0.67 | 0.91 | 5.2 | 0.60 | 0.45 | 0.45 | 0.60 | 620 |
| linkedin.com | 7.8 | 1.13 | 0.93 | 0.91 | 0.84 | 5.2 | 0.75 | 0.62 | 0.61 | 0.56 | 15 |
| chatgpt.com | 7.8 | 1.31 | 0.76 | 0.49 | 0.69 | 5.2 | 0.87 | 0.51 | 0.33 | 0.46 | 18 |
| instagram.com | 7.8 | 1.56 | 0.77 | 0.57 | 0.70 | 5.2 | 1.04 | 0.52 | 0.38 | 0.47 | 96 |
| x.com | 7.8 | 1.76 | 1.15 | 1.70 | 1.74 | 5.2 | 1.17 | 0.77 | 1.14 | 1.16 | 144 |
| en.wikipedia.org | 7.8 | 0.91 | 1.13 | 1.13 | 1.09 | 5.2 | 0.60 | 0.75 | 0.75 | 0.72 | 583 |
| dadrian.io/blog/... | 7.8 | 1.82 | 1.50 | 1.50 | 1.52 | 5.2 | 1.21 | 1.00 | 1.00 | 1.01 | 8 |
| cnn.com | 7.8 | 2.29 | 0.87 | 0.96 | 1.04 | 5.2 | 1.53 | 0.58 | 0.64 | 0.70 | 6,120 |
| msn.com | 7.8 | 1.40 | 1.01 | 1.72 | 1.05 | 5.2 | 0.93 | 0.67 | 1.15 | 0.70 | 614 |
| stackoverflow.com | 7.8 | 1.78 | 0.99 | 0.92 | 1.33 | 5.2 | 1.18 | 0.66 | 0.61 | 0.89 | 33 |
| tiktok.com | 7.8 | 1.53 | 1.25 | 0.43 | 0.81 | 5.2 | 1.02 | 0.83 | 0.29 | 0.54 | 73 |

Table 1: Calculated Standard ML-DSA (15KB cert chain) vs. Trimmed ML-DSA (10KB) web metric slowdowns on 4G Networks (9Mbps down/9Mbps up, 170ms RTT) with cached browser content (data collected from `webpagetest.org`. We can see that although MLWE's impact on the handshake is almost 8%, the web metrics will notice only 1-2%.




Figure 2: Numerous parallel connections opened for `cnn.com` (experimental results from `webpagetest.org`).

Contentful Paint (FCP), and Document Complete (DC) times and showed it will be much less significant than on the handshake itself. Table 1 shows some of these results. The study also demonstrated that popular websites open multiple TLS 1.3 connections which are not sequential, and, as such, even if they transfer small amounts of data and are rendered on the page, they will not have cumulative impact on the webpage performance. Figure 2 shows some of these connections for `cnn.com`. What is more, [6] looked into some popular websites and confirmed that the web content includes multiple small connections that collect web metrics, track users, or include HTTP cached content, redirects etc, but the vast size of web content rendered on browsers comes from just a few connections of significant size. These findings seem to align with Cloudflare's data in [23].

This new study builds on the investigations in [6] by measuring actual web performance metrics. We explore the impact of the MLWE algorithms with example web content served from five different servers in

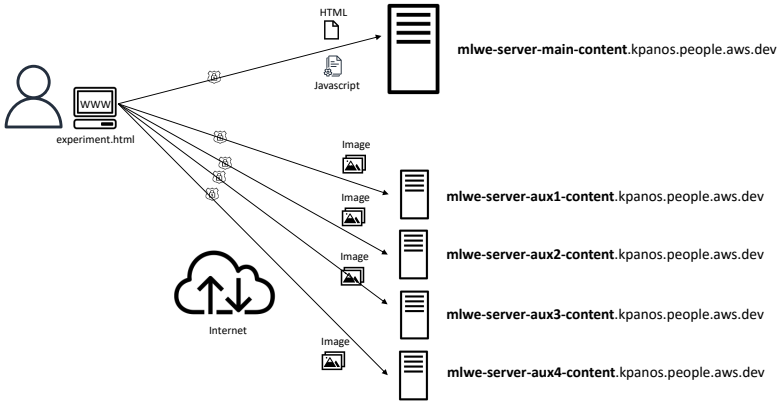



Figure 3: The experimental setup.

`HTML`, simple `Javascript`, and images. To experiment with real-world web scenarios, some of this content is cacheable (HTTP **Etag**) so it does not need to be transferred every time. The images are fetched in parallel uncorrelated connections, but we also experiment with sequential fetches as aa worst case scenario. Figure 3 shows the architecture of the webpage. The content is served by `nginx` and `OpenSSL 3.5`. The metrics data is collected with `webpagetest.org`.

To get an extra intuition, we also compare the example webpage's performance with TLS 1.2 and TLS 1.3. TLS 1.2 includes an extra round-trip per connection which can contribute slowdowns per handshake that are worse than the transmission delay of the 15-20KB from ML-KEM and ML-DSA. For example, at 5Mbps, a typical 30ms round-trip is equivalent to the transfer time of 18.75KB.

We also share data about TLS 1.3 connections in common AWS Services like which use RESTful APIs to showcase the point that many of these use-cases will not suffer significantly as also argued previously in [8].

The main contributions of this work are

1. It explores real-world web scenarios with small and larger, cached or non-cached web content and quantifies (with actual measurements) qualitative results from previous studies [8, 6] about the new algorithms' impact on web metrics.
2. It quantifies (with actual measurements) the impact of one additional round-trip to real-world web content scenarios since one round-trip could be similar to the transmission delay of the additional MLWE data introduced in these connections.
3. It goes beyond the typical web scenario and showcases that other common TLS use-cases in the cloud will not see significant impact.

# References


[1] C. Bas Westerbaan. Sizing Up Post-Quantum Signatures, Nov. 2021.

[2] D. Benjamin, D. O'Brien, and B. Westerbaan. Merkle Tree Certificates for TLS. Internet-Draft draft-davidben-tls-merkle-tree-certs-01, Internet Engineering Task Force, Sept. 2023. Work in Progress.

[3] K. Bürstinghaus-Steinbach, C. Krauß, R. Niederhagen, and M. Schneider. Post-quantum tls on embedded systems: Integrating and evaluating kyber and sphincs+ with mbed tls. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, ASIA CCS '20, page 841–852, New York, NY, USA, 2020. Association for Computing Machinery.

[4] T. Iraklis, K. Limniotis, and N. Kolokotronis. Evaluating the performance of post-quantum secure algorithms in the TLS protocol. *Journal of Surveillance, Security and Safety*, 3(3):101–27, 2022.

[5] D. Jackson. Abridged Compression for WebPKI Certificates. Internet-Draft draft-ietf-tls-cert-abridge-02, Internet Engineering Task Force, Sept. 2024. Work in Progress.

[6] P. Kampanakis. How much will ML-DSA Signatures affect Web Metrics after all? `https://pkic.org/events/2025/pqc-conference-austin-us/THU_BREAKOUT_1130_Panos-Kampanakis_How-much-will-ML-DSA-affect-Webpage-Metrics.pdf`, 2025.

[7] P. Kampanakis, C. Bytheway, B. Westerbaan, and M. Thomson. Suppressing CA Certificates in TLS 1.3. Internet-Draft draft-kampanakis-tls-scas-latest-03, Internet Engineering Task Force, Jan. 2023. Work in Progress.

[8] P. Kampanakis and W. Childs-Klein. The impact of data-heavy, post-quantum TLS 1.3 on the Time-To-Last-Byte of real-world connections, 2024.

[9] P. Kampanakis and M. Kallitsis. Faster post-quantum TLS handshakes without intermediate CA certificates. In S. Dolev, J. Katz, and A. Meisels, editors, *Cyber Security, Cryptology, and Machine Learning*, pages 337–355, Cham, 2022. Springer International Publishing.

[10] P. Kampanakis and T. Lepoint. Vision paper: Do we need to change some things? In F. Günther and J. Hesse, editors, *Security Standardisation Research*, pages 78–102, Cham, 2023. Springer Nature Switzerland.

[11] L. V. Kris Kwiatkowski. The TLS Post-Quantum Experiment, Oct. 2020.

[12] A. Langley. CECPQ1 results, Nov. 2016.

[13] A. Langley. CECPQ2, Dec. 2018.

[14] D. Marchsreiter and J. Sepúlveda. Hybrid post-quantum enhanced tls 1.3 on embedded devices. In *2022 25th Euromicro Conference on Digital System Design (DSD)*, pages 905–912, 2022.

[15] NIST. Migration to Post-Quantum Cryptography.

[16] NIST NCCOE PQC Migration Project. NIST SPECIAL PUBLICATION 1800-38C Draft.

[17] C. Paquin, D. Stebila, and G. Tamvada. Benchmarking post-quantum cryptography in tls. In J. Ding and J.-P. Tillich, editors, *Post-Quantum Cryptography*, pages 72–91, Cham, 2020. Springer International Publishing.

[18] S. Paul, Y. Kuzovkova, N. Lahr, and R. Niederhagen. Mixed certificate chains for the transition to post-quantum authentication in tls 1.3. In *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security*, ASIA CCS '22, page 727–740, New York, NY, USA, 2022. Association for Computing Machinery.

[19] E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446, Aug. 2018.

[20] D. Sikeridis, S. Huntley, D. Ott, and M. Devetsikiotis. Intermediate certificate suppression in post-quantum tls: An approximate membership querying approach. In *Proceedings of the 18th International Conference on Emerging Networking EXperiments and Technologies*, CoNEXT '22, page 35–42, New York, NY, USA, 2022. Association for Computing Machinery.

[21] D. Sikeridis, P. Kampanakis, and M. Devetsikiotis. Assessing the overhead of post-quantum cryptography in TLS 1.3 and SSH. In *Proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies*, CoNEXT '20, page 149â€“156, New York, NY, USA, 2020. Association for Computing Machinery.

[22] D. Sikeridis, P. Kampanakis, and M. Devetsikiotis. Post-quantum Authentication in TLS 1.3: A performance study. In *27th Annual Network and Distributed System Security Symposium, NDSS 2020, San Diego, California, USA, February 23-26, 2020*. The Internet Society, 2020.

[23] B. Westerbaan and L. Valenta. A look at the latest post-quantum signature standardization candidates. `https://blog.cloudflare.com/another-look-at-pq-signatures/`, Nov. 2024.