

Post-Quantum Authentication at Netflix

Ian McDiarmid-Sterling, Srikanth Varadarajan, Hoss Shafagh, Jared Crawford

Netflix

September 2, 2026

Executive Summary

As part of Netflix's mission to entertain the world, we are migrating to post-quantum cryptography (PQC) to ensure the continued security and privacy of Netflix data. Our goal is to make this transition seamless for our members, allowing them to continue watching their favorite content uninterrupted. Netflix's migration to PQC focuses on defending against two main threats:

1. Harvest-now-decrypt-later attacks, in which attackers collect long-lived encrypted data today to decrypt it later with a quantum computer.
2. Forge-later attacks, where attackers use quantum computing to forge authentication in systems that rely on long-term trust.

This document will primarily address our efforts to mitigate forge-later attacks by deploying post-quantum (PQ) authentication.

The expected lifetime of some of the devices Netflix members use already intersects with established PQC migration timelines. For these devices, we are collaborating with device ecosystem partners to prioritize the transition to PQ authentication to mitigate forge-later attacks. This collaboration ensures that we continue to provide a cryptographic foundation for bootstrapping both the device and the Netflix application. Similarly, we are prioritizing changes to the Netflix backend and core infrastructure to support establishing secure communication channels from Netflix devices using PQ authentication.

We are intentionally deferring our PQC migration for assets not impacted by harvest-now-decrypt-later or forge-later attacks and instead prioritizing further investment in crypto agility. This investment in crypto agility will allow us to adopt post-quantum algorithms and parameters efficiently over time.

Member Devices

Netflix members can interact with content on many types of devices including smart TVs, set-top boxes, streaming sticks, mobile phones, personal computers, game consoles, and more. On these devices, we typically need to bootstrap software trust (e.g., secure boot, secure software updates) and subsequently bootstrap secure communications with the Netflix backend. You can think of this fundamental distinction as the difference between trust that the device is running legitimate, signed firmware and trust that the device is talking to Netflix's legitimate backend. At Netflix, keys used for both trust models are difficult to rotate.

This means that the keys used to bootstrap software trust must provide adequate security over the device's lifetime. For our longer-lived (7+ years) devices, the device lifetime will intersect with PQC migration timelines (est. by 2030-2035). This intersection is one of the key drivers behind our overall strategy, motivating us to bootstrap software trust via hybrid algorithms that combine classical and post-quantum primitives. As we engage with device ecosystem partners, one of the primary challenges we face is addressing concerns around increased key and signature sizes as the industry continues to face memory (e.g. DRAM) and storage (e.g. NAND Flash, eMMC) constraints.

Protocols

The switch to PQ authentication in our protocols is not a substitution problem; we are spending significant time reducing the number of required signatures and preparing for future algorithmic upgrades. In cases where we don't need non-repudiation and the threat model permits it, we prefer transitioning impacted protocols from signature-based authentication to authentication based on key derivation, using either hash-based message authentication codes (HMAC) or authenticated encryption with associated data (AEAD) schemes. This is not a clean replacement; it trades one set of guarantees for another, but allows us to use more efficient algorithms.

Hardware Security Modules (HSMs)

Many of our cryptographic protocols are eventually terminated in FIPS-140-3 Level 3 validated HSMs. This means that as we adopt PQ authentication in our device ecosystem, we simultaneously require support for PQ authentication algorithms in our HSM fleet. As we actively collaborate with our preferred cloud providers to shape HSM roadmaps, we are simultaneously preparing for the lower signing performance of ML-DSA compared to our current ECC workloads. This eventual performance degradation raises two key challenges for us:

- 1) How do we ascertain the n-th order impact to our microservices of slower authentication operations in our HSM fleet?
- 2) How do we establish the right business tradeoffs to minimize the financial impact of Netflix's cryptographic protocols while protecting the member experience?

To tackle the first challenge, we have built on Netflix's existing failure injection system, designing the ability to selectively inject latency on requests without impacting our members. This allows us to benchmark the impact of slower HSM operations to all microservices in the call paths, before the underlying HSM implementation is released. This strategy of experimentally validating any assumptions as early as possible is a key strategic bet we are making to reduce the amount of uncertainty in our PQC migration – allowing us to intentionally defer work, scope more accurately, and avoid faulty assumptions about how our microservices behave.

To tackle the second challenge, we are iterating on our HSM capacity management system so that we can run a larger fleet (to support PQ authentication operations) while minimizing added cost. Traffic to our HSM fleet varies dramatically; daily viewing patterns produce peaks and troughs, and live

events/show launches lead to large traffic spikes. More efficient capacity prediction algorithms allow us to capitalize on this variance and effectively offset the cost of switching to PQ authentication.

Crypto Agility

Over the last decade, we have invested in centralized tooling (a.k.a paved roads) for core cryptography that serves a variety of enterprise use cases (encrypting credit card data, protecting personally identifiable information, etc). This has allowed us to maintain an inventory of cryptographic usage throughout the business, provided us with a single point of leverage for algorithmic changes, and minimized implementation risk. We believe that in an era of AI-assisted cryptanalysis, crypto agility is our most durable hedge against algorithmic weakening/breakages, whether classical or post-quantum.

As we look ahead, we are reaffirming this strategic choice by actively expanding paved support for additional cryptographic usage, such as token-based (JOSE) cryptography. To fully automate these flows, we are publishing reference vectors, building hooks to unblock the rotation of keys used externally, and establishing requirements for token-based verifiers (internal and external). While many of Netflix's token-based use cases are actively blocked from upgrading to PQ authentication (ex. partners have limited algorithmic support, tokens are size-constrained, etc.) we are making the intentional choice not to dedicate resources to unblocking these efforts since our token-based cryptography is short-lived and not impacted by harvest-now-decrypt-later/forge-later attacks. By focusing on centralization rather than immediate migration, we give blockers time to resolve naturally while optimally positioning ourselves for when they do.

Conclusion

Netflix's commitment to securely entertaining the world requires us to prioritize support for post-quantum authentication in our device ecosystem, protocols, and HSM fleet while simultaneously preparing for a broader, enterprise-wide migration to post-quantum cryptography. While we face ongoing challenges addressing the impact of larger public keys, reducing the cost of HSM-based ML-DSA, and encouraging vendors to support PQ authentication, our focus on crypto agility and experimentation will allow us to continue delivering a seamless member experience.