Move Fast and Roll Your Own Crypto

A Quick Look at the Confidentiality of Zoom Meetings

By Bill Marczak and John Scott-Railton April 3, 2020

Download this report

Read our description of Zoom's waiting room vulnerability, as well as frequently asked question about Zoom and encryption issues.

This report examines the encryption that protects meetings in the popular Zoom teleconference app. We find that Zoom has "rolled their own" encryption scheme, which has significant weaknesses. In addition, we identify potential areas of concern in Zoom's infrastructure, including observing the transmission of meeting encryption keys through China.

Key Findings

- Zoom documentation claims that the app uses "AES-256" encryption for meetings where possible. However, we find that in each Zoom meeting, a single AES-128 key is used in ECB mode by all participants to encrypt and decrypt audio and video. The use of ECB mode is not recommended because patterns present in the plaintext are preserved during encryption.
- The AES-128 keys, which we verified are sufficient to decrypt Zoom packets intercepted in Internet traffic, appear to be generated by Zoom servers, and in some cases, are delivered to participants in a Zoom meeting through servers in China, even when all meeting participants, and the Zoom subscriber's company, are outside of China.
- Zoom, a Silicon Valley-based company, appears to own three companies in China through which at least 700 employees are paid to develop Zoom's software. This arrangement is ostensibly an effort at *labor arbitrage*: Zoom can avoid paying US wages while selling to US customers, thus increasing their profit margin. However, this arrangement may make Zoom responsive to pressure from Chinese authorities.

1. Background: A US Company with a Chinese Heart?

Zoom is a popular teleconference app whose popularity has increased dramatically, given much of the world is under mandatory work-from-home orders due to the spread of COVID-19. The app's overarching design goal seems to be reducing friction in videoconferencing and making things "just work."

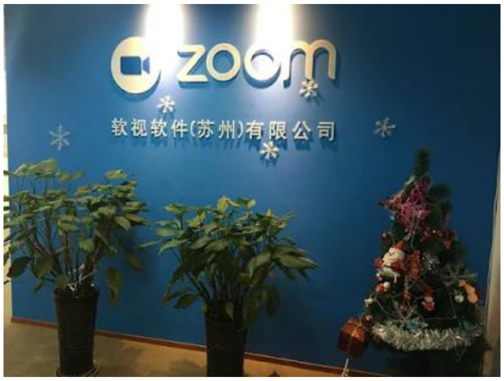


Figure 1: A picture shows the Zoom logo above the name of one of Zoom's Chinese developer companies, "Ruanshi Software (Suzhou) Ltd." (Source)

While Zoom is headquartered in the United States, and listed on the NASDAQ, the mainline Zoom app appears to be developed by three companies in China, which all have the name 软视软件 ("Ruanshi Software"). Two of the three companies are owned by Zoom, whereas one is owned by an entity called 美国云视频软件技术有限公司 ("American Cloud Video Software Technology Co., Ltd.") Job postings for Ruanshi Software in Suzhou include open positions for C++ coders, Android and iOS app developers, and testing engineers.

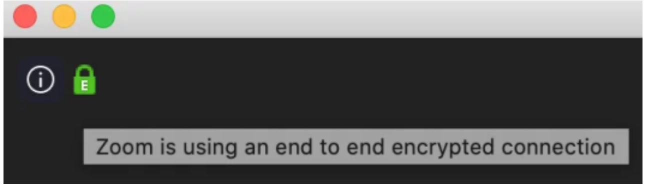
Zoom's most recent SEC filing shows that the company (through its Chinese affiliates) employs at least 700 employees in China that work in "research and development." The filing also implies that 81% of Zoom's revenue comes from North America. Running development out of China likely saves Zoom having to pay Silicon Valley salaries, reducing their expenses and increasing their profit margin. However, this arrangement could also open up Zoom to pressure from Chinese authorities. While the mainline Zoom app (zoom.us) was reportedly blocked in China in November 2019, there are several third-party Chinese companies that sell the Zoom app within China (e.g., zoom.cn, zoomvip.cn, zoomcloud.cn).

Any Feature You Like, As Long As It's Speed

In the past few years, a number of security issues regarding Zoom have come to light. These issues have included unintentional bugs, such as vulnerabilities in Zoom's screen sharing feature, and privacy concerns, such as Zoom sharing data with Facebook. However, perhaps the most prominent security issues with Zoom surround deliberate features designed to reduce friction in meetings, which also, by design, reduce privacy or security. This includes Zoom installing a hidden web-server on Mac computers to circumvent a Safari popup that users had to click through before they joined a Zoom meeting, a Zoom feature that removes a password prompt during the installation process (and instead displays a misleading password prompt later), a Zoom feature intended to allow Zoom users at the same company (or ISP) to easily find each other, and Zoom's easy 9 or 10 digit code which is sufficient to join a meeting created with default settings, leading to the well-reported phenomenon of "Zoom Bombing."

Encryption Questions Come to Light

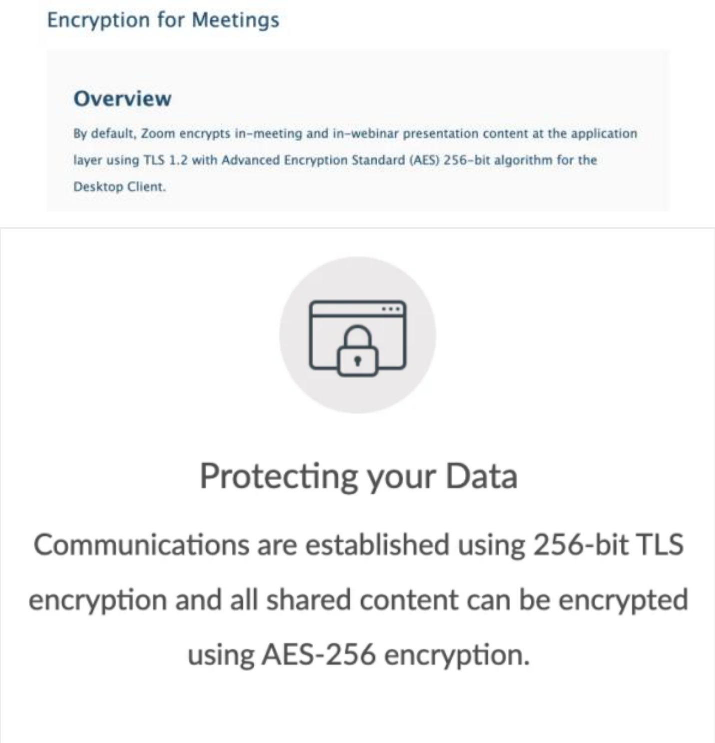
Zoom's documentation has a number of unclear claims about encryption that the platform offers. Some Zoom documentation (as well as the Zoom app itself) claims that Zoom offers a feature for "end-to-end (E2E) encrypted meetings.



Zoom is using an end to end encrypted connection

Figure 2: Zoom's app displays a message incorrectly claiming that a call is "end-to-end" encrypted.

Typically, the computer security community understands the term "end-to-end encrypted" to mean that *only* the parties to the communication can access it (and not any middlemen that relay the communication). Other Zoom documentation says that Zoom's meeting software for Windows, MacOS, and Linux "by default" uses the industry-standard TLS 1.2 scheme for transport encryption, though a September 2014 blog post implies that this software does *not* use TLS.



Encryption for Meetings

Overview

By default, Zoom encrypts in-meeting and in-webinar presentation content at the application layer using TLS 1.2 with Advanced Encryption Standard (AES) 256-bit algorithm for the Desktop Client.

Protecting your Data

Communications are established using 256-bit TLS encryption and all shared content can be encrypted using AES-256 encryption.

Figure 3a and 3b: Zoom claims regarding TLS and AES encryption (Source: Zoom documentation, Zoom website).

In response to this confusion, Zoom released a blog post in April 2020 describing their encryption scheme. The blog post clarifies that Zoom does not currently implement "end-to-end" encryption as most people understand the term; Zoom used the term "end-to-end" to describe a situation where all conference participants (except those dialing in via the public switched telephone network) are required to use *transport* encryption between their devices and Zoom servers. Zoom's definition of "end-to-end" does not seem to be a standard one, even in the realm of enterprise videoconferencing solutions. Because Zoom does not implement true end-to-end encryption, they have the theoretical ability to decrypt and monitor Zoom calls. Nevertheless, Zoom mentions that they have not built any mechanism to intercept their customers meetings: *"Zoom has never built a mechanism to decrypt live meetings for lawful intercept purposes, nor do we have means to insert our employees or others into meetings without being reflected in the participant list."*

Zoom's April 2020 blog post does not, however, provide details about exactly how their encryption works, or clarify whether they use TLS or AES-256. Because of the potentially misleading and conflicting claims regarding Zoom's encryption, and the proliferation of Zoom's technology in the business, government, civil society, and healthcare sectors where confidentiality may be desired, we decided to examine exactly how Zoom meetings are encrypted.

2. COVID-19: A New Gold Rush for Cyber Spies

Social distancing and work-from-home policies have shifted government, economic, and personal activity online. In the rush to reconnect, users are rapidly adopting new apps and communications platforms. Some popular video chat and collaboration tools have added millions of users, almost overnight. In many cases, consumer choice appears to be driven by the need for usability, speed, and stability, rather than careful assessment of privacy policies and security protocols.

At the same time, the newly remote workforce is heavily reliant on personal equipment and online accounts for work business. The shift away from work networks and accounts denies cyber defenders the ability to enforce security standards, while blocking their visibility into potential compromises.

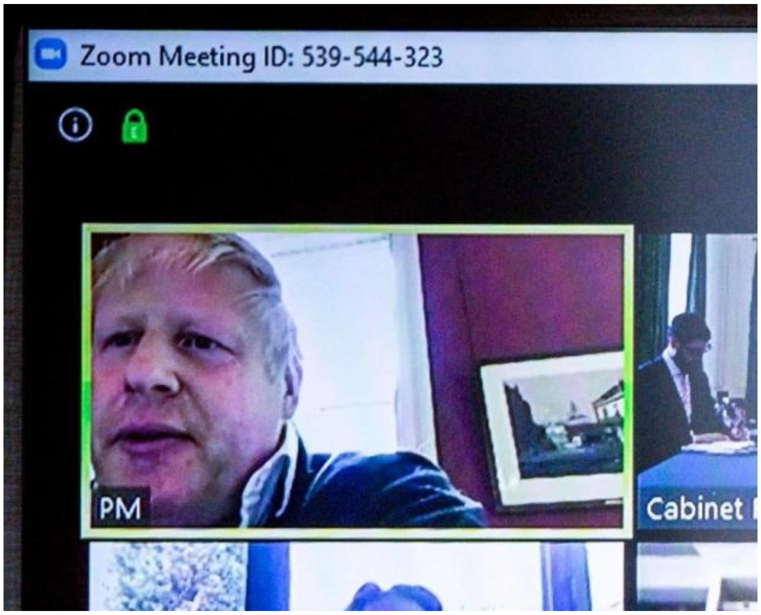


Figure 4: UK PM Boris Johnson conducting a cabinet meeting over Zoom (Source).

Interactions that were previously conducted in the real world are now mediated by popular digital platforms. Until a few weeks ago, it would have been uncommon for high stakes business negotiations, high level diplomacy, political strategy conferences, and cabinet meetings to be conducted over platforms whose security properties are unknown. Eavesdropping on these encounters would have been out of reach to all but the most sophisticated digital adversaries.

Now, some of the most sensitive conversations in the world are taking place on devices and platforms vulnerable to basic forms of eavesdropping and attack techniques. This “new normal” is a potential goldmine for cyber spies. Given the business value of meetings currently being conducted on Zoom, it is reasonable to expect that the platform is being closely scrutinized by groups engaged in industrial and political espionage, and cybercrime.

Zoom as an Intelligence Target

Zoom’s success has led it to attract conversations that are of high priority interest to multiple governments. We suspect that this makes Zoom a high priority target for signals intelligence (SIGINT) gathering and targeted intrusion operations.

Most governments conduct electronic espionage operations. Their targets include other governments, businesses, and individuals. Some, including the Chinese government, are known to conduct extensive industrial espionage. In addition, a growing number of governments have sought out mobile phone hacking technology and abused it to target the personal phones of journalists, lawyers, judges, and others who seek to hold them to account.

In addition, as digital rights advocacy group Access Now has pointed out in an open letter calling for a transparency report, Zoom has not publicly disclosed information such as statistics of requests for data by governments, and what Zoom has done in response to these requests. Zoom’s policies concerning notifications to users over breaches or the handing-over of data to governments are also unknown, however the company has just promised at the time of writing to release such a report within 90 days of April 2nd.

3. Results: Custom Crypto, Chinese Servers, Security Issues

Rather than using a standard protocol for sending voice and video, Zoom appears to implement their own transport protocol. The Zoom transport protocol appears to be a bespoke extension of the existing RTP standard.

The Zoom transport protocol adds Zoom’s own encryption scheme to RTP in an unusual way. By default, all participants’ audio and video in a Zoom meeting appears to be encrypted and decrypted with a single AES-128 key shared amongst the participants. The AES key appears to be generated and distributed to the meeting’s participants by Zoom servers. Zoom’s encryption and decryption use AES in ECB mode, which is well-understood to be a bad idea, because this mode of encryption preserves patterns in the input. Industry standard protocols for encryption of streaming media (e.g., the SRTP standard) recommend the use of AES in Segmented Integer Counter Mode or f8-mode, which do not have the same weakness as ECB mode. **Figure 5** is a classic illustration of the perils of ECB mode: the outline of a penguin is still visible in an image encrypted with ECB mode.[1]

Figure 5: A classic illustration of why ECB mode is not recommended. An image of a penguin (left) is encrypted in ECB mode and then visualized (right). Note that the outline of the penguin remains visible in the encrypted image (Source: Wikipedia).

During a test of a Zoom meeting with two users, one in the United States and one in Canada, we found that the AES-128 key for conference encryption and decryption was sent to one of the participants over TLS from a Zoom server apparently located in Beijing, 52.81.151.250. A scan shows a total of five servers in China and 68 in the United States that apparently run the same Zoom server software as the Beijing server. We suspect that keys may be distributed through these servers. A company primarily catering to North American clients that sometimes distributes encryption keys through servers in China is potentially concerning, given that Zoom may be legally obligated to disclose these keys to authorities in China.

Figure 6: The topology of our Zoom test call.

During our analysis, we also identified a security issue with Zoom's Waiting Room feature. Assessing that the issue presented a risk to users, we have initiated a responsible vulnerability disclosure process with Zoom. We are not currently providing public information about the issue to prevent it from being abused. We intend to publish details of the vulnerability once Zoom has had a chance to address the issue. In the meantime, Section 5 provides recommendations on how users can mitigate the issue.

Figure 7: A Zoom call participant in the Waiting Room, to be admitted to a Meeting Room.

4. How we Investigated

We began by observing Internet traffic associated with Zoom meetings using the Zoom clients on Windows, MacOS, and Linux. We used Wireshark to record our Internet traffic while we joined and participated in Zoom meetings. The vast majority of the Internet traffic during our Zoom meetings was exchanged between our computer and servers owned by Zoom on UDP port 8801. A further examination of the UDP traffic revealed that Zoom had apparently designed their own transport protocol, which wraps the well-known RTP protocol for transferring audio and video.

Identifying Encrypted Video

On some packets, whose UDP payload began with 0x05100100, the RTP header often encoded a type value of 98. In these packets, the RTP payload appeared to contain an H.264 video stream using the format in RFC 6184. In this format, the RTP payload is a series of one or more NALUs (Network Abstraction Layer Units), which carry components of the video (e.g., various types of video frames, metadata on decoder settings, etc). Some of the NALUs were *fragmented* using the scheme from the RFC for "Fragmentation Unit A" (FU-A). We re-assembled these into unfragmented NALUs. Per the RFC, each NALU has a "type value" indicating which component of the video it carries. In Zoom's case, all of the NALU values were set to zero, which is invalid per the RFC, so we suspected that the NALU payload was a format bespoke to Zoom.

Each NALU payload consisted of a 4-byte big-endian value that appeared to describe a length (these 4-byte values were all less than, but close to the size of the packets), followed by a number of bytes that was always the lowest multiple of 16 larger than the 4-byte length value (i.e., if the 4-byte length value was between 145 and 160, it would be followed by 160 bytes). This suggested to us the use of the AES encryption scheme, which operates on blocks of 16 bytes. If the length of a message to be encrypted is not a multiple of 16 bytes, then *padding* is added to the end of the message to inflate the length to a multiple of 16. An examination of a memory dump of the Zoom process during a meeting revealed an AES-128 key in memory associated with the string *conf.skey*, which we speculated stood for "conference secret key."

Figure 8: A novelty placard and a Citizen Lab notebook are visible in this frame of H.264 video we extracted from a PCAP of a Zoom call and decrypted using the AES-128 conf.skey in ECB mode.

To extract video for each participant, we first grouped the RTP packets by the SSRC (Synchronization Source Identifier) value in the RTP header. Each SSRC value indicates a single participant. For each SSRC, we reassembled fragmented H264 NALUs in the correct order using RTP timestamps and sequence numbers, then decrypted them with the AES-128 key in ECB mode, then de-padded the decrypted result (using the 4-byte length value), and finally wrote the decrypted data to disk in a raw H.264 stream file. We were able to play the file using the following VLC media player command:

```
$ vlc raw.h264 --demux h264
```

Identifying Encrypted Audio

We also noticed other packets in our Wireshark capture that began with the header value 0x050f0100 and the RTP header in these packets often contained a type value of 112. In these packets, the RTP timestamp was incremented by 640 between subsequent packets. We located a research paper that describes how to infer the type of RTP audio codecs by looking at various metadata in the RTP packets, including the difference between the RTP timestamps in subsequent packets. The paper provides one possibility for a timestamp difference of 640, which is the Skype-developed SILK codec, at a 16000Hz sample rate. We also noted that the RTP payloads in these packets appeared to have a similar encryption format as the NALU payloads in the video packets, though they appeared to contain a two-byte rather than four-byte length header.

To extract audio for each participant, we first grouped the RTP packets by SSRC. For each participant (SSRC), we created a SILK file, beginning with the magic bytes "#!SILK_V3". For each SSRC, we decrypted the bytes following the two-byte length value (using the same AES-128 key in ECB mode). We wrote the decrypted bytes, prepended with the two-byte length value (in little-endian byte order) from the RTP payload. We then obtained a SILK transcoder and successfully transcoded each SILK file into an MP3 containing the audio from one of the participants.

```
$ sh converter.sh raw.silk mp3
```

Figure 9 shows the layers of encapsulation involved in both Zoom video and audio packets.

Figure 9: A protocol layer diagram showing the encapsulation present on Zoom video (left) and audio (right) packets.

Identifying AES Key Transmission

We next sought out to discover how the meeting's AES-128 encryption key (*conf.skey*) was derived. We noticed that before the large amount of traffic on UDP port 8801, there was some TLS traffic between our computer and Zoom servers. We set up mitmproxy to intercept the TLS traffic and configured the Zoom Linux client to route its TLS traffic through mitmproxy.[2] Fortunately, the Zoom client *did* appear to warn us that the fake TLS certificates generated by mitmproxy were untrusted. After we trusted the certificates, we observed a series of messages exchanged between our Zoom client and Zoom servers. In one of the messages, the Zoom server sent us the encryption key in **Figure 10**.

Figure 10: An example of an AES-128 conf.skey transmitted from the Zoom server to our Zoom client and decrypted with mitmproxy.

It is unclear to us whether Zoom servers use a cryptographically secure random number generator to create the meeting encryption keys or whether the keys may somehow be predictable. We confirmed that all participants in a Zoom meeting have the same *conf.skey* value and that this key does *not* change when participants join or leave. The key does, however, change when *all* users leave the meeting for a period of time; any new participant joining an empty meeting will cause the generation of a new *conf.skey* value.

5. Conclusion: Not Suited for Secrets

Zoom's product is user-friendly and has rapidly grown its user base during the COVID-19 pandemic by "just working." Zoom's fast growing user base, combined with marketing language around encryption and security, have attracted many sensitive conversations. This sudden popularity likely puts the product in the crosshairs of government intelligence agencies and cybercriminals.

Questionable Crypto & Encryption Keys Sent to Beijing

Unfortunately for those hoping for privacy, the implementation of call security in Zoom may not match its exceptional usability. We determined that the Zoom app uses non-industry-standard cryptographic techniques with identifiable weaknesses. In addition, during multiple test calls in North America, we observed keys for encrypting and decrypting meetings transmitted to servers in Beijing, China.

An app with easily-identifiable limitations in cryptography, security issues, and offshore servers located in China which handle meeting keys presents a clear target to reasonably well-resourced nation state attackers, including the People's Republic of China.

Our report comes amidst a number of other recent research findings and lawsuits identifying other potential security and privacy concerns with the Zoom app. In addition, advocacy groups have also pointed out that Zoom lacks a transparency report, a critical step towards addressing concerns arising when companies have access to sensitive user data. Zoom has just stated (April 2nd, 2020) that it will release such a report within 90 days.

As a result of these troubling security issues, we discourage the use of Zoom at this time for use cases that require strong privacy and confidentiality, including:

- Governments worried about espionage
- Businesses concerned about cybercrime and industrial espionage
- Healthcare providers handling sensitive patient information
- Activists, lawyers, and journalists working on sensitive topics

For those using Zoom to keep in touch with friends, hold social events, or organize courses or lectures that they might otherwise hold in a public or semi-public venue, our findings should not necessarily be concerning.

For those who have no choice but to use Zoom, including in contexts where secrets may be shared, we speculate that the browser plugin may have some marginally better security properties, as data transmission occurs over

TLS.

Use Zoom Passwords, Avoid Waiting Rooms

As part of our research, we identified what we believe to be a serious security issue with Zoom's Waiting Room feature. We have initiated a responsible disclosure process with Zoom, which is currently being responsive. We hope that the company will quickly act to patch and provide an advisory. In the meantime, we advise Zoom users who desire confidentiality to ***not* use Zoom Waiting Rooms**. Instead, we encourage users to **use Zoom's password feature**, which appears to offer a higher level of confidentiality than waiting rooms. Instructions on password features can be found here.

Scrutiny Needed

The rapid uptake of teleconference platforms such as Zoom, without proper vetting, potentially puts trade secrets, state secrets, and human rights defenders at risk. Companies and individuals might erroneously assume that because a company is publicly listed or is a major household name, that this means the app is designed using security best practices.

As we showed in this report, that assumption is false.

Acknowledgements

Thanks to Masashi Nishihata, Miles Kenyon, and Lotus Ruan.

Bill Marczak's work on this report is partially supported by the Center for Long Term Cybersecurity (CLTC) at UC Berkeley and the International Computer Science Institute,

The Citizen Lab is grateful for support from the Ford Foundation and the John D. and Catherine T MacArthur Foundation.

1. [1] Note that the penguin image on the left of Figure 5 is an uncompressed bitmap. If it were compressed (e.g., a JPEG or PNG), visualizing the outline of the encrypted penguin on the right would be somewhat more difficult.↩
2. [2] The Zoom Linux client allows a feature for explicitly configuring an HTTPS proxy, whereas the Mac and Windows clients do not appear to have this feature. ↩