Push Away Your Privacy: Precise User Tracking Based on TLS Client Certificate Authentication

Matthias Wachs, Quirin Scheitle, and Georg Carle
TMA’17, Dublin, June 21, 2017
TLS 1.2 handshake does not encrypt certificates

*Known for a long time...*

![TLS 1.2 handshake diagram](image)

**Server Certificates**
- Eavesdroppers can learn the specific websites that a user visits (not just the server’s IP address)

**Client Certificates**
- Used by VPNs, governments, ...
- Person names, company names, ... → private data!
TLS 1.2 Client Certificate Authentication (CCA)

Where is CCA used?

- **Network authentication**: 802.1x EAP
- **VPN**: OpenVPN, F5 EdgeConnect, …
- **Web**: HTTPS
- **IoT**: MQTT
- **Remote device management**, for example MobileIron
- **Apple Push Notification Service (APNs)**

Apple Statistics:

- 1 billion active devices (2016)
- 800 million iTunes accounts (2014)
Push Notification Services

Figure: Push Service Architecture: Messages brokered to Apps through the Push Notification Service.

Resource efficient notification of (mobile) applications:

- **Apple’s APNs**: OS, MacOs, iTunes
- **Google’s FCM**: Android, Chrome
- **Microsoft’s WNS**: Windows, Windows Phone

Paradigms:

- Tightly integrated with operating system
- Always connected to backend
Apple Push Notification Service (APNs)

*Maybe the biggest user of TLS CCA?*

APNs integral part of iOS and macOS – “always on”

APNs uses Client Certificates for login:
- Generated at device setup
- Unique cryptographic material (CN, public key, fingerprint)
- CN different for mobile and desktop devices

Serial Number: ab:12:34:56:78:9a:bc:de:f0:12
Issuer: C=US, O=Apple Inc., OU=Apple iPhone, CN=Apple iPhone Device CA
Validity Not Before: Apr 8 12:34:56 2015 GMT
Validity Not After: Apr 8 12:34:56 2016 GMT
Subject: CN=12345678-1234-1234-1234-123456789ABC
Key ...
(all data redacted)
Precise User$^1$ Tracking in APNs

Several appearances of same device easily linkable

2 of 4 Attacker Types Considered in this Work

- Apple or someone infiltrating Apple: better means available
- Local adversary: Can use MAC addresses and more
- Regional adversary: Access to one or several large networks
- Global adversary: Access to several core networks

Regional Adversary – Validation at Internet Uplink

- Can a regional adversary track users?

Global Adversary – Validation through Global Path Measurements

- How well can a global adversary leverage APNs to track users?

1: APNs CCA certificates are bound to devices. However, these devices are typically private and carried by a user at most times, which allows inferences into user tracking.
Passive Capturing

Methodology

Analysis of > 2 weeks of TLS CCA traffic at Internet uplink:

- APNs TCP ports (443, 5223, 2195, 2196)
- \texttt{pcap} Filter on certificate handshake

Stored information:

- Timestamp
- Connection 5-tupel (Source & Destination IP address, Port, Protocol TCP)
- Certificates & TLS Extensions
Working with Human Subjects

Ethical Considerations

Strict regulations by IRB:

- Documented measurement process
- Isolated measurement infrastructure
- Access only for permitted staff
- Raw data must not leave infrastructure

Our self-restrictions:

- No attempt to identify users
- No publication of identifiable data
APNs by far the biggest user of CCA

<table>
<thead>
<tr>
<th>#Certs</th>
<th>Issuer Distinguished Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>56128</td>
<td>/C=US/O=Apple Inc./OU=Apple iPhone/CN=Apple iPhone Device CA</td>
</tr>
<tr>
<td>334</td>
<td>/CN=Layer Client CA/C=US/L=San Francisco/O=Layer, Inc/ST=CA</td>
</tr>
<tr>
<td>221</td>
<td>/CN=AnyDesk Client</td>
</tr>
<tr>
<td>76</td>
<td>/C=KR/ST=Kyunggido/L=Suwon/O=Samsung Electronics (redacted)</td>
</tr>
<tr>
<td>52</td>
<td>/CN=Ricoh Remote Service (redacted)</td>
</tr>
</tbody>
</table>
Case Study - how well can we track a single user?

Informed Consent

Note: We are tracking a device. As mobile devices are typically closely carried, they allow conclusions about users.
What % of certificates is traceable?

![Graph showing the percentage of certificates seen on n days for iOS APNs Certificates, All APNs Certificates, and Desktop APNs Certificates.]

Quirin Scheitle (TUM) | Push Away Your Privacy: Precise User Tracking Based on TLS CCA
Can we derive device types from certificate data?

% of certificates with 'valid before' date

- iOS APNs Certificates
- Desktop APNs Certificates

Vertical Lines:
- Apple Product Launches
- iOS certificate expiry
- Desktop certificate expiry

Date
Is global tracking feasible?

**Methodology**

Research Question: How many networks does an attacker have to eavesdrop on to observe a significant share of APNs logins?

- We identify APNs backend infrastructure and conduct distributed traceroute measurements towards it
  - Measurements confirm that clients resolve one of `[1-50]-courier.push.apple.com`
  - We globally resolve `[1-50]-courier.push.apple.com` using 1000 RIPE Atlas probes each
  - We find 69 /24 subnets and pick one random observed IP address in each of the 69 subnets
  - Using 1000 RIPE Atlas probes per measurement, we conduct traceroute measurements towards all 69 IP addresses
- We map transit router’s IP addresses to ISPs and IXPs
- We count what % of routes traverses a certain ISP or IXP
Is global tracking feasible?

_Eavesdropping capabilities on just 10 networks allows to follow APNs messages of over 80% of users globally or nationally_

<table>
<thead>
<tr>
<th>Rank</th>
<th>Global</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IXP/AS</td>
<td>Σ% Paths</td>
</tr>
<tr>
<td>1</td>
<td>AS3356 (L3)</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>AS1299 (Telia)</td>
<td>40%</td>
</tr>
<tr>
<td>3</td>
<td>AS174 (Cogent)</td>
<td>54%</td>
</tr>
<tr>
<td>4</td>
<td>AS7922 (Comcast)</td>
<td>61%</td>
</tr>
<tr>
<td>5</td>
<td>AS12322 (Free)</td>
<td>67%</td>
</tr>
<tr>
<td>6</td>
<td>AS6830 (Liberty)</td>
<td>71%</td>
</tr>
<tr>
<td>7</td>
<td>AS4637 (Telstra)</td>
<td>75%</td>
</tr>
<tr>
<td>8</td>
<td>AS6453 (Tata)</td>
<td>78%</td>
</tr>
<tr>
<td>9</td>
<td>AS2828 (XO)</td>
<td>81%</td>
</tr>
<tr>
<td>10</td>
<td>AS3320 (DTAG)</td>
<td>84%</td>
</tr>
</tbody>
</table>

Note: % is based on RIPE Atlas probe distribution as a proxy for APNs user distribution.
Responsible Disclosure

We informed Apple’s product security team before publication:

• Contact with OpenPGP secured mail
• Very quick response
• Several phone calls, continuous contact
• Several engineers in calls and working on resolution

Impact:

• MacOS & iOS fixed with January 2017 security patches
• APNs Backend patched
• iTunes on Windows patched a bit later (SChannel is complicated . . . )
It has been known and criticized for a while that TLS1.2 does not encrypt certificates, which may have specific adverse impact for client certificates. Anyhow . . .

• Discussions eroded . . .
• Draft RFCs expired . . .
• Apple decided to use CCA for APNs . . .

Lack of taking the issue seriously?

We believe that Internet measurements can overcome inertia in security improvements by . . .

• Quantifying impact and scale of a problem with hard evidence
• Benefitting issue prioritization
• Providing means to track patching progress
What now?

*Push TLS 1.3 standardization which encrypts certificates*

---

**Figure: TLS 1.3 handshake, Unencrypted Data, [Encrypted Data]**

But: **ClientHello Extensions still unencrypted:**

- Server Name Indication (SNI)
- Application-specific data
Reproducibility.

We aim for repeatability, replicability, and reproducibility\(^1\)

Repeatability — same team, same experimental setup
Packing of “Reproducibility Bundle” along with camera-ready version requires detailed repetition of paper creation.

Replicability — different team, same experimental setup
We provide “artifacts” (scripts, data, documentation) so any other team can easily replicate our work.

Reproducibility — different team, different experimental setup
We provide a detailed documentation of our approach (\textit{which pcap filter was set? what precise traceroute parameters were set?}) so other teams can reproduce our work without using our artifacts.

We ran an exercise\(^2\) at the TMA PhD school that followed the research question and methodology of this paper. This resulted in a partial mix of replication, reproduction, and extension of our work.

---

1: Terms as defined by ACM: http://www.acm.org/publications/policies/artifact-review-badging
2: https://github.com/quirins/tma17-ripeatlas-lab-participants/
How to deal with Reproducibility and Private Data?

*Much of the data in this work contains private and sensitive data*

### Passively Captured TLS handshakes and certificates

- No publication of raw data
- Cut open of analysis pipeline (for example, “not valid before” attribute of certificate)
- Anonymize output of database query with **documented** script
- Feed the anonymized data into analysis pipeline, published – figures in paper clickable: https://github.com/tumi8/cca-privacy/blob/master/userstudy/userstudy.ipynb

### Active Measurement of APNs backend and traceroutes

- Public data – RIPE Atlas measurements per default public, for example
  https://atlas.ripe.net/measurements/5719601/
- Publish everything: Measurement scripts, RIPE Atlas IDs, raw data, analysis tools
Future Work

- Measuring uptake of APNs patch
- In-depth analysis of APNs backend infrastructure
- Controlling for AS population vs. RIPE Atlas probe count bias
Key Messages, Data, and Code

- TLS-CCA sends certificates unencrypted
- In an “always-on” mobile scenario, this can cause serious privacy issues
- We quantified this issue in the Apple Push Notification Service (APNs), Apple fixed promptly

Data and Code:
https://github.com/tumi8/cca-privacy

Matthias Wachs, **Quirin Scheitle**, and Georg Carle
Chair of Network Architectures and Services — https://net.in.tum.de
Technical University of Munich (TUM)