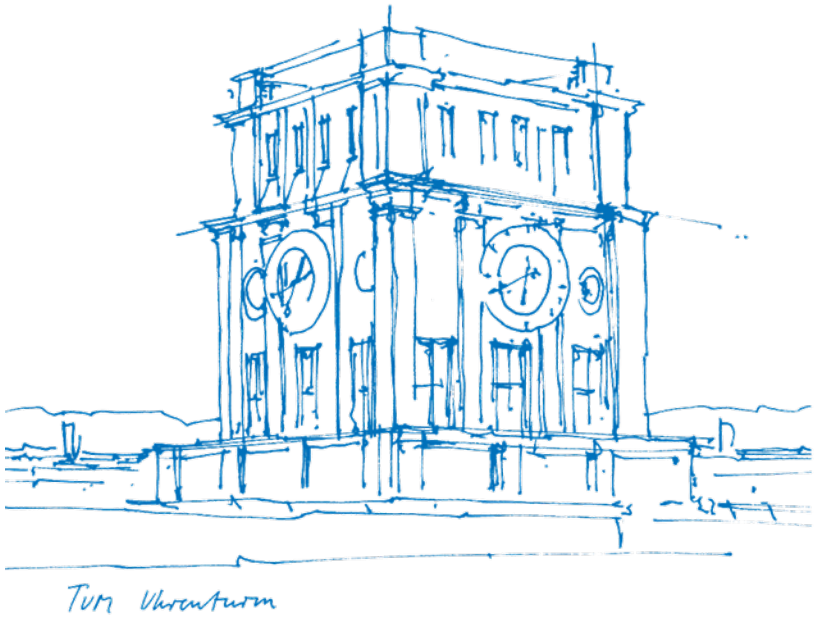
Chair of Network Architectures and Services
TUM Department of Informatics
Technical University of Munich (TUM)

TMA 2017 PhD School — RIPE Atlas Lab

Emile Aben, Quirin Scheitle

Dublin, June 20, 2017



Agenda

Learning Goals

- Become familiar with RIPE Atlas and its capabilities and use
- Run DNS and Traceroute measurements using RIPE Atlas
- See how to plan measurements and analyze data along the lines of a TMA'17 paper

Outline

- Background on Research Questions
- DNS Measurements
- Traceroute Measurements
- Mapping to ASes/IXPs

Background

Apple Push Notification Service (APNs)

Maybe the biggest user of unencrypted TLS Client Certificate Authentication?

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

APNs uses Client Certificates for login:

- Generated at device setup
- Very unique cryptographic material (CN, public key, fingerprint)

```
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 Tracking in APNs

Several appearances of same device easily linkable

2 attacker types

- ~~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 – Feasibility 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? **This exercise**

Detailed Research Questions and Approach

Research Question: How many networks do you need to eavesdrop on to surveil a majority of APNs backend logins?

Steps to be taken:

- What are the backend servers?
- How to measure user population connecting to backend servers?
- How to map this to networks? What is a network in this context?

Step 1: Finding APNs Backend Servers

From passive observations at 1 Vantage Point (VP), we know that clients resolve [1-50]-courier.push.apple.com and then connect to 1 IP address in the 17.0.0.0/8 range.

How to find all APNs backend servers?

Step 1: Finding APNs Backend Servers

From passive observations at 1 Vantage Point (VP), we know that clients resolve [1-50]-courier.push.apple.com and then connect to 1 IP address in the 17.0.0.0/8 range.

How to find all APNs backend servers?

Task: Pick 1 of ‘[1-50]-courier.push.apple.com‘ and do a RIPE Atlas DNS resolution for your home country

Example: 42-courier.push.apple.com

Note: More experienced Ripe Atlas users are invited to do resolutions for all 50 DNS names.

Step 0: Prerequisites

- RIPE Atlas Account
- RIPE Atlas Credits
- Python3
- Github Download:

```
git clone --recursive git://github.com/quirins/tma17-ripeatlas-lab-participants
```

Step 1: Finding APNs Backend Servers - Discussion

- What probes did you choose? How many per country?
- Which detailed DNS settings did you choose?
- How did you run the measurement? How to scale it to 50?

New DNS measurement: `https://atlas.ripe.net/measurements/form/`
Sample measurement from paper: `https://atlas.ripe.net/measurements/5500016/`
Sample measurement from June 2017: `https://atlas.ripe.net/measurements/8831682`
Script for batch measurements:
`https://github.com/tumi8/cca-privacy/blob/master/ripe_atlas/dns/atlas-measure.sh`

Step 1: Finding APNs Backend Servers - Obtaining the Result

Your measurement should have finished by now – please obtain the result and parse it

Our parsing script:

`https://github.com/tumi8/cca-privacy/blob/master/ripe_atlas/dns/parse-results.py`

Step 1: Finding APNs Backend Servers - Obtaining the Result

Your measurement should have finished by now – please obtain the result and parse it

Our parsing script:

`https://github.com/tumi8/cca-privacy/blob/master/ripe_atlas/dns/parse-results.py`

Discussion

- Download via Browser or REST
- JSON with abuf
- Region-specific CNAMEs
- Are these all IP addresses? Or just some?

Sample measurement result from June 2017: `https://github.com/quirins/tma17-ripeatlas-lab-participants/blob/master/data/dns/RIPE-Atlas-measurement-8831682.json`
Sample parsed result from paper `https://github.com/quirins/tma17-ripeatlas-lab-participants/blob/master/data/dns/result-5500014.json.parsed.txt`
Sample parsed result from June 2017 `https://github.com/quirins/tma17-ripeatlas-lab-participants/blob/master/data/dns/RIPE-Atlas-measurement-8831682.json.parsed.txt`

Step 2: Traceroute Path Measurements

Our DNS queries have yielded a list of backend servers. Coming back to our Research Question, we want to quantify the number of networks an adversary has to eavesdrop on to see a significant number of logins directed to those backend servers.

Task: Define and execute a measurement strategy: Which RIPE Atlas settings? Which probes? Which targets?

Note: Again, Ripe Atlas novices can just run measurements towards 1 target IP address.

Step 2: Traceroute Path Measurements

Our DNS queries have yielded a list of backend servers. Coming back to our Research Question, we want to quantify the number of networks an adversary has to eavesdrop on to see a significant number of logins directed to those backend servers.

Task: Define and execute a measurement strategy: Which RIPE Atlas settings? Which probes? Which targets?

Note: Again, Ripe Atlas novices can just run measurements towards 1 target IP address.

Discussion

- Traceroute Details – `https://atlas.ripe.net/measurements/form/`
- Target Selection – all IP addresses? Some?
- Probe Selection – which probes to select? Do these represent the APNs user base
  - Which probes to select?
  - Do they represent the APNs user base? AS/CC bias?
  - Only probes that resolved the IP being probed?

Traceroute measurement from Paper - Germany: `https://atlas.ripe.net/measurements/5719601/`

Step 2: Traceroute Path Measurements

Task: Please download the results – what format does it have? Ideas how to parse it? What would be the next steps?

Step 2: Traceroute Path Measurements

Task: Please download the results – what format does it have? Ideas how to parse it? What would be the next steps?

- JSON – Ripe Atlas Cousteau or raw JSON parsing
- Next step: Map to IXPs and ASes – Ideas on data sources?

Step 2: Traceroute Path Measurements

Task: Please download the results – what format does it have? Ideas how to parse it? What would be the next steps?

- JSON – Ripe Atlas Cousteau or raw JSON parsing
- Next step: Map to IXPs and ASes – Ideas on data sources?

Fortunately, the link `https://github.com/tumi8/cca-privacy/tree/master/ripe_atlas/traceroute` contains the script “traceroutes_to_asn_ixp.py.”

Task: Call it using ./traceroutes_to_asn_ixp.py your-measurement.json ip2ixp ip2as
TODO: Link Intermediate results

Step 3: Analyze Results

How does the result look like for your country? Is that in line with the paper?

Rank	Global		Germany	
	IXP/AS	Σ% Paths	IXP/AS	Σ% Paths
1	AS3356 (L3)	25%	IXP DE-CIX	30%
2	AS1299 (Telia)	40%	AS3320 (DTAG)	52%
3	AS174 (Cogent)	54%	IXP E-CIX	61%
4	AS7922 (Comcast)	61%	AS6830 (Liberty)	69%
5	AS12322 (Free)	67%	AS31334 (VF/Kabel D)	75%
6	AS6830 (Liberty)	71%	AS1273 (C&W)	78%
7	AS4637 (Telstra)	75%	AS3356 (L3)	81%
8	AS6453 (Tata)	78%	AS34419 (VF Group)	84%
9	AS2828 (XO)	81%	AS680 (DFN)	86%
10	AS3320 (DTAG)	84%	AS6805 (Telefonica)	88%

Step 3: Analyze Results

How does the result look like for your country? Is that in line with the paper?

Rank	Global		Germany	
	IXP/AS	Σ% Paths	IXP/AS	Σ% Paths
1	AS3356 (L3)	25%	IXP DE-CIX	30%
2	AS1299 (Telia)	40%	AS3320 (DTAG)	52%
3	AS174 (Cogent)	54%	IXP E-CIX	61%
4	AS7922 (Comcast)	61%	AS6830 (Liberty)	69%
5	AS12322 (Free)	67%	AS31334 (VF/Kabel D)	75%
6	AS6830 (Liberty)	71%	AS1273 (C&W)	78%
7	AS4637 (Telstra)	75%	AS3356 (L3)	81%
8	AS6453 (Tata)	78%	AS34419 (VF Group)	84%
9	AS2828 (XO)	81%	AS680 (DFN)	86%
10	AS3320 (DTAG)	84%	AS6805 (Telefonica)	88%

What do you think of the result overall?

Feedback & Conclusion

- How did you like the exercise?
- Useful parts / less useful parts?
- Excited to use Ripe Atlas? :)

OUT / Backup

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

Rank	Global		Germany	
	IXP/AS	Σ% Paths	IXP/AS	Σ% Paths
1	AS3356 (L3)	25%	IXP DE-CIX	30%
2	AS1299 (Telia)	40%	AS3320 (DTAG)	52%
3	AS174 (Cogent)	54%	IXP E-CIX	61%
4	AS7922 (Comcast)	61%	AS6830 (Liberty)	69%
5	AS12322 (Free)	67%	AS31334 (VF/Kabel D)	75%
6	AS6830 (Liberty)	71%	AS1273 (C&W)	78%
7	AS4637 (Telstra)	75%	AS3356 (L3)	81%
8	AS6453 (Tata)	78%	AS34419 (VF Group)	84%
9	AS2828 (XO)	81%	AS680 (DFN)	86%
10	AS3320 (DTAG)	84%	AS6805 (Telefonica)	88%

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

Data, and Code

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`
TUM Informatics