iTAP: In-network Traffic Analysis Prevention using Software-Defined Networks

Roland Meier, David Gugelmann, Laurent Vanbever

https://itap.ethz.ch

N.S.A. May Have Hit Internet Companies at a Weak Spot

The Internet companies’ data centers are locked down with full-time security [...] But *between the data centers* [...] information was unencrypted and an easier target for government intercept efforts, according to three people with knowledge of Google’s and Yahoo’s systems who spoke on the condition of anonymity.


---

Google encrypts data amid backlash against NSA spying

*By Craig Timberg*  September 5, 2013

Google is racing to encrypt the torrents of information that flow among its data centers around the world in a bid to stave off a new wave of foreign governments’ sweeping surveillance programs that have been making life difficult for the Web giant.

The move by Google and other leading Internet firms that recently revealed they had been conduits for the National Security Agency’s sweeping surveillance program is a response to demands from American technology companies, Internet advocates, and various legal authorities.
Existing solutions

Do not protect communicating parties
[SSL/TLS, IPsec Transport, MACsec]

Require modifications at end-hosts or additional middleboxes
[APOD, CONTRA]

Do not support partial deployment or have scalability problems
[MACsec, PHEAR]

More references provided in the paper
In–network Traffic Analysis Prevention using Software–Defined Networks

Roland Meier, David Gugelmann, Laurent Vanbever
In-network Traffic Analysis Prevention using Software-Defined Networks
In-network Traffic Analysis Prevention using Software-Defined Networks

- Communication anonymity
  who is communicating with whom?
In-network Traffic Analysis Prevention using Software-Defined Networks

- Communication anonymity: who is communicating with whom?
- Volume anonymity: how much traffic flows between X and Y?
iTAP

In-network Traffic Analysis Prevention using Software-Defined Networks

- Communication anonymity
  who is communicating with whom?

- Volume anonymity
  how much traffic flows between X and Y?

- Topology anonymity
  how many hosts are in the network?
In-network Traffic Analysis Prevention using Software-Defined Networks

- No modifications at end-hosts
iTAP

In-network Traffic Analysis Prevention using Software-Defined Networks

- Central controller
- Rewriting capabilities of switches
An iTAP–protected network
An iTAP-protected network

Layer 2 network
An iTAP-protected network

Layer 2 network

With some SDN switches
An iTAP–protected network

Layer 2 network
With some SDN switches
And a central controller
An iTAP–protected network

- Layer 2 network
- With some SDN switches
- And a central controller
- Attacked by an eavesdropper
An iTAP-protected network

Layer 2 network
With some SDN switches
And a central controller
Attacked by an eavesdropper
Protected by iTAP
Example
Example
Packet from A to B enters the network
Ingress switch notifies controller
Controller computes & installs flow rules
Ingress switch obfuscates source and destination
Core switch forwards obfuscated packet

Controller

A

B

A → B

2ρη

2ρη
Egress switch de-obfuscates source and destination
How does the rewriting work?
Rewriting packet headers

Trade–off between anonymity and scalability

iTAP approach: Mixing per–host IDs and random bits

Measure information leakage & counteract attacker
Rewriting packet headers as a trade–off between anonymity and scalability
Rewriting packet headers as a trade-off between anonymity and scalability

- Unique ID per flow
Rewriting packet headers as a trade-off between anonymity and scalability

- Unique ID per flow
- Unique ID per host
Rewriting packet headers as a trade-off between anonymity and scalability

- Unique ID per flow
- iTAP hybrid approach
- Unique ID per host
iTAP hybrid obfuscation scheme
iTAP hybrid obfuscation scheme
iTAP hybrid obfuscation scheme

Map source and destination to IDs

01001001  A  B  00110111
iTAP hybrid obfuscation scheme

Map source and destination to IDs

Match-fields with arbitrary bitmasks

<table>
<thead>
<tr>
<th>MAC src</th>
<th>MAC dst</th>
<th>IP src</th>
<th>IP dst</th>
</tr>
</thead>
<tbody>
<tr>
<td>01001001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00110111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
iTAP hybrid obfuscation scheme

Map source and destination to IDs

Match-fields with arbitrary bitmasks

Interpret as bit-string of 160 bits

<table>
<thead>
<tr>
<th>MAC src</th>
<th>MAC dst</th>
<th>IP src</th>
<th>IP dst</th>
</tr>
</thead>
<tbody>
<tr>
<td>01001001</td>
<td>00110111</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
iTAP hybrid obfuscation scheme

Map source and destination to IDs

Match-fields with arbitrary bitmasks

Interpret as bit-string of 160 bits

Randomly select bits that are used for source and destination ID
**iTAP hybrid obfuscation scheme**

Map source and destination to IDs

<table>
<thead>
<tr>
<th>MAC src</th>
<th>MAC dst</th>
<th>IP src</th>
<th>IP dst</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Match-fields with arbitrary bitmasks

Interpret as bit-string of 160 bits

Randomly select bits that are used for source and destination ID

Add **source** and **destination** ID

| 0 0 0 1 1 1 0 0 0 1 0 1 0 1 1 1 |
iTAP hybrid obfuscation scheme

Map source and destination to IDs

Match-fields with arbitrary bitmasks
Interpret as bit-string of 160 bits
Randomly select bits that are used for source and destination ID
Add source and destination ID
Set other bits to random values
iTAP hybrid obfuscation scheme

Map source and destination to IDs

Match-fields with arbitrary bitmasks

Interpret as bit-string of 160 bits

Randomly select bits that are used for source and destination ID

Add source and destination ID

Set other bits to random values
iTAP hybrid obfuscation scheme

Forwarding based on the destination ID → good scalability
iTAP hybrid obfuscation scheme

Eavesdropper cannot distinguish between **random** and non-random bits
\rightarrow good anonymity
What if an attacker analyzes multiple flows?
What if an attacker analyzes multiple flows?
What if an attacker analyzes multiple flows?
iTAP controls information leakage and proactively adapts the encoding

The controller monitors the observed entropy for each link...

... and changes the encoding before an eavesdropper is able to break it.
iTAP controls information leakage and proactively adapts the encoding

The controller monitors the observed entropy for each link...

... and changes the encoding before an eavesdropper is able to break it.*

* According to the Unicity Distance
iTAP controls information leakage and proactively adapts the encoding
iTAP controls information leakage and proactively adapts the encoding
iTAP controls information leakage and proactively adapts the encoding
iTAP controls information leakage and proactively adapts the encoding.
iTAP evaluation based on real network traffic

iTAP evaluation based on

7 days of network traffic
400 hosts
128 M flows
iTAP evaluation based on real network traffic

7 days of network traffic  400 hosts  128 M flows

Indicators:
controller actions / s
flow table updates / s
forwarding rules
iTAP works in practice

7 days of network traffic 400 hosts 128 M flows

avg max
200 700 controller actions / s
50 250 flow table updates / s
600 2.5 k forwarding rules
Only a small share of SDN switches is sufficient to protect a large share of the network traffic.

<table>
<thead>
<tr>
<th>Protection level</th>
<th>100% –</th>
</tr>
</thead>
</table>

- Linear topology
- Tree topology

SDN-enabled switches
Only a small share of SDN switches is sufficient to protect a large share of the network traffic.
Only a small share of SDN switches is sufficient to protect a large share of the network traffic.
Contributions

iTAP design

Scalable & anonymity-providing header rewriting scheme

iTAP prototype implementation

Evaluation based on real user traffic

https://itap.ethz.ch