Middlebox Measurement and Cooperation

Mirja Kühlewind and Brian Trammell, ETH Zürich

CleanSky Workshop
Heidelberg, 29 Feb 2016
Problem Statement:
Ossification of the Internet due to Middlebox Impairments

Problem
Middleboxes make restrictive, implicit assumptions about traffic passing through them

➡ Deployment of "new" protocols/extension limited by packet/flow modifications of middleboxes

Goal
Reduce the accidental manipulation to zero, while minimizing the essential manipulation!

Needed
1. More data about the nature and distribution of middlebox impairments
   ➡ Common data model for storage and analysis of middlebox impairment

2. Explicit Middlebox cooperation to declare assumptions and intentions independent of the used transport or higher-layer protocol
   ➡ New (UDP-based) transport encapsulation + in-band signaling
The MAMI Project
Measurement and Architecture for a Middleboxed Internet

- Strong interaction with relevant standards organizations for impact on deployment
- FIRE testbed (MONROE) support for measurement as well as experimentation, especially on mobile broadband access networks
- Learn more at http://mami-project.eu/
Middlebox Measurements:
Golas and Overview

1. Large-scale measurements of path impairments
   - using FIRE MONROE as well as RIPE Atlas, CAIDA Ark…
   - UDP/TCP/SCTP connectivity, TCP options (e.g. TFO, MPTCP), and other protocol (ICMP, DNS, …)

2. Development of new measurements tools: https://github.com/mami-project/
   - Tracebox: tracing + impairment analysis
   - PathSpider: A/B testing (currently on ECN support)

3. Path Transparency Observatory
   - Active measurements by the project + external measurements
   - Query interface to access observations on path impairments:
     - What is the likelihood that a certain path impairment impacts my traffic (modifications/stripping/dropping/blocking)?
Path Transparency Observatory

- Observatory (public release end 2016) to derive common observations about conditions on a given path at a given time
- Combining disparate measurements leads to better insight
  - e.g. own measurement data, traceroutes, BGP, traces

Follow [http://mami-project.eu](http://mami-project.eu) for availability!
Is it possible to run the Internet over UDP? Preliminary Results

- A/B testing for TCP/UDP connectivity
- Copycat tool on 120 PlanetLab nodes
  - 3.67% UDP blocking on port 33435
  - 2.7% UDP blocking on all tested ports (33435, 1228, 8008, 12345)
- RIPE Atlas traceroute
  - 3.661% UDP blocking based on existing traceroutes
- We are currently running more measurements!
- Use all existing testbeds available, e.g. CAIDA Ark, MONROE
Middlebox Cooperation: Architectural Considerations

1. Shim for Middlebox Cooperation Protocol (MCP)
   - Transport and applications can selectively expose semantic information to middlebox
   - Higher layers can fully be encrypted

2. Flexible Transport Layer (FTL)
   - Maintain connectivity (even if the MCP is not supported) e.g. fallback or happy-eyeball mechanisms
   - Provision of encryption context for different layers/protocols
Why a new shim?

- **Transport layer**: end-to-end sockets
  - flow information
  - stateful and ‘smart’ processing at the edge
- **Internet layer**: hop-by-hop handling
  - per-packet information
  - stateless and simple processing in the middle
Why a new shim?

- Transport layer: end-to-end sockets
- flow information
- stateful and complex processing at the edge
- stateless and simple processing in the middle

➡ Path layer for explicit cooperation with middleboxes instead of implicit assumptions

Missing: Per-flow information for stateful in-network functions

architecture
Path Layer: (Basic) Functional Requirements

- Grouping of packets into flows
- Extensibility to provide per-flow network information
- Explicit feedback channel

<table>
<thead>
<tr>
<th>magic</th>
</tr>
</thead>
<tbody>
<tr>
<td>tube/group/flow id</td>
</tr>
<tr>
<td>resv</td>
</tr>
<tr>
<td>option space …</td>
</tr>
<tr>
<td>checksum</td>
</tr>
</tbody>
</table>
Example 1: Firewall Traversal

Problem

UDP often blocked as it is hard to maintain state

Needed

- group ID
- start/stop signal and confirmation by receiver (’SYN/ACK’)

Action

- firewall can forward first packet and set up state based on confirmation from receiver
- group ID must be large enough to not be guessable
Example 2: Low Latency Support

Problem

Network service not optimized for latency sensitive traffic

Needed

Flag to signal loss sensitivity vs. latency sensitivity

Action

• network device can treat latency sensitive traffic differently, e.g. in a separate smaller queue
• trade-off between loss and latency gives not incentive to lie
Why should I trust what you say about your flows?

- Default: *trust but verify*
  - declarative signaling: **no** negotiation, **no** guarantees
  - the best way to prevent cheating is to make it useless to do so

- Leverage existing trust relationships for higher-assurance declarations
  - e.g. your enterprise firewall, access network middleboxes, etc.
References

• Substrate Protocol for User Datagrams (SPUD) in the IETF
  • draft-trammell-spud-req
  • draft-kuehlewind-spud-use-cases
  • draft-hildebrand-spud-prototype
• IAB Stack Evolution Program
  • Workshop on Stack Evolution in a Middlebox Internet (SEMI) 2015 [RFC7663]
  • B. Trammell, J. Hildebrand: Evolving Transport in the Internet
• IRTF proposed research group on Measurement and Analysis for Protocols (MAPRG)
• MAMI webpage (mami-project.eu) or twitter (@mamiproject)
Summary and Conclusion

Problem
Ossification of the Internet Protocol Stack

Needed
1. Measurement to identify path impairments
   - Large-scale using all available testbeds (incl. MONROE)
   - New measurements tools (Tracebox, PathSpider)
   - Path Transparency Observatory
2. Path layer for explicit middlebox cooperation
   - Middlebox Cooperation Protocol (MCP): trust by verify
   - Encrypted everything else!