Enabling Renewed Innovation in TCP by Establishing an Isolation Boundary

Umar Kalim\textsuperscript{*,+}, Eric Brown\textsuperscript{+,}, Mark K. Gardner\textsuperscript{+,}, Wu-chun Feng\textsuperscript{*}

Department of Computer Science\textsuperscript{*}, Office of IT\textsuperscript{+}
Virginia Tech
11/28/2010
Why Renewed Innovation?

• Desirable functionality
  – Multipath TCP / Multihoming
  – Flow Migration
Why Renewed Innovation?

• Desirable functionality
  – Hybrid Transports

Motivation

Motivation

- **Existing solutions**
  - TCP option space
    - MPTCP addresses the issue
  - Custom libraries
    - Duplication of effort
  - Clean-slate approach

- **Legacy behind TCP**
  - Experience and tools built to support TCP
  - Best to adopt an incremental approach
Proposed Solution

• Goals
  – Extract commonality in current work
  – Lay the foundation for higher layer services
  – Admit incremental adoption

• Common mechanism
  – Decouple application stream from transport flows
  – Construct a control channel

• Leverage TCP options
Outline

• Motivation
• Proposed Solution
  – Isolation Boundary: An Overview
  – Connection Setup
  – Control and Data Option
• Community Engagement
• Contribution and Implications
Isolation Boundary: An Overview

- Decouple application-data stream from transport-endpoint identification
  - Avoid use of 4-tuple for transport endpoint ID
- TCP Isolation Boundary Option
- Partially adapt Isolation Layer proposed by Next Generation Transport (Tng)*

Transport Independent Flow Setup

- Along with TCP Handshake

**Peer** subscripts A and B are used.

**Choose TIFID**

Choose TIFID\(_A\) and TISeq\(_A\),

\[
\text{SYN} + \text{TIFID}_A + \text{TISeq}_A
\]

**Record TIFID**

Record TIFID\(_B\), TISeq\(_B\),

\[
\text{SYN} + \text{ACK} + \text{TIAck}_A + \text{TIFID}_B + \text{TISeq}_B
\]

\[
\text{ACK} + \text{TIAck}_B
\]

**Interference? Fallback!**

**Transport Independent Flow ID (TIFID)**

<table>
<thead>
<tr>
<th>Sequence #</th>
<th>Acknowledgement #</th>
</tr>
</thead>
</table>

TISeq – Transport Independent Sequence #
TIAck – Transport Independent Acknowledgement #

Virginia Tech
Invent the Future
Isolation Boundary Options

- Isolation Boundary Option - Control
  - Admits out-of-band control channel
- Control protocol to be defined by community
Isolation Boundary Options

- Not all applications need a control channel
- Isolation Boundary Options
  - Control
  - Data
Isolation Boundary Options

- Isolation boundary enables reconnection
  - From network fault

Reuse TIFID and TISeq<sub>A</sub>

Ack TISeq<sub>B</sub>

Connection Established

Disconnection

Identify TIFID, Ack. TISeq<sub>A</sub>

ACK + TIAck<sub>B</sub>
Flow Termination

- Control Channel to clear state
- Exchange of TCP FINs implies cleanup
- Rely on timeouts for network faults
Community Engagement

- Standardization
  - Specification of the extensible control channel
  - Reference implementation
- Wire Protocol Specification
- Isolation Boundary for other transports
Contribution and Implications

• Develop an Isolation Boundary
  – Decouple the entity naming from transport-endpoint identification
  – Construct a control channel to facilitate higher layer services

• Infrastructure support for
  – Hybrid transport
  – Multihoming
  – Flow migration over networks
  – Flow migration between processes
  – Reconnection after network fault
  – ... and other innovations as the community thinks of them
Thank you

• Contact
  – Umar Kalim - umar@cs.vt.edu
  – Eric Brown - brownej@vt.edu
  – Mark Gardner - mkg@vt.edu
  – Wu-chun Feng – feng@cs.vt.edu

• SyNeRGy
  – http://synergy.cs.vt.edu

This research is supported in part by Juniper Networks and Virginia Tech
Analysis

• TCP Options Space
  – 3-way handshake ~ 20 octets available
    • MSS, Window Scaling, SACK, Timestamp

• Incompatible Options
  – Alternate checksum, Partial Ordering, Transactional TCP, TCP MD5, TCP Authentication, Quick Start Response

• Performance
  – Exchange is off the critical data path (3-way handshake)

• SYN Cookies
  – IBO not preserved when under attack
Analysis

• Middleboxes
  – Fall back to legacy TCP if options are stripped

• Security
  – No worse than TCP
  – To hijack a session the attacker must know:
    • Transport Independent Flow ID
      – Exchanged only during 3-way handshake
    • Sequence numbers for unacknowledged data
      – Each TCP segment must be accounted for, to derive current state from Initial Seq. Nos.

• Application Compatibility
  – Backward compatible